

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:17:45 ; Search time 19 Seconds  
(without alignments)  
233.835 Million cell updates/sec

Title: US-09-868-953A-2

Perfect score: 807

Sequence: 1 MFSTKALLLAGLSTALGP.....GIQSHVHINCSDPTAYL 151

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents.AA.\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep.\*

2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep.\*

3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep.\*

4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep.\*

5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep.\*

6: /cgn2\_6/ptodata/1/1aa/backfilest1.pep.\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	806	99.9	151	2	US-08-563-148E-6
2	802	99.4	151	4	US-09-194-139-1
3	645	79.9	133	1	US-07-924-753-15
4	421	52.2	98	4	US-09-194-139-9
5	256	31.7	54	2	US-08-563-148E-1
6	196	24.3	41	1	US-07-924-753-1
7	132	16.4	27	1	US-07-924-753-14
8	95	11.8	20	2	US-08-563-148E-5
9	94	11.6	18	1	US-07-924-753-4
10	90	11.2	19	1	US-07-924-753-3
11	89	11.0	19	1	US-07-924-753-12
12	86	10.7	18	1	US-07-924-753-11
13	74	9.2	702	6	5240838-15
14	73	9.0	14	1	US-07-924-753-10
15	70.5	8.7	625	3	US-08-996-139-15
16	70.5	8.7	625	4	US-08-995-659-15
17	70.5	8.7	625	4	US-09-215-649A-15
18	70.5	8.7	625	4	US-09-577-780-15
19	70	8.7	10182	4	US-09-134-001C-3159
20	69	8.6	14	1	US-07-924-753-9
21	67.5	8.4	2285	4	US-09-308-375-2
22	67	8.3	14	4	US-09-194-139-10
23	66	8.2	2004	1	US-08-375-709-15
24	66	8.2	2004	1	US-08-752-929-15
25	66	8.2	2004	4	US-09-090-793-9
26	65.5	8.1	406	4	US-08-293-778-24
27	65.5	8.1	406	1	US-08-295-411-5

28	65.5	8.1	406	2	US-08-955-471-5	Sequence 5, Appli
29	65.5	8.1	406	5	PCT-US92-10242-5	Sequence 5, Appli
30	65.5	8.1	444	1	US-08-475-845-2	Sequence 2, Appli
31	65.5	8.1	444	2	US-08-327-690-2	Sequence 2, Appli
32	65.5	8.1	444	2	US-08-660-289-2	Sequence 2, Appli
33	65.5	8.1	444	2	US-08-537-807-2	Sequence 2, Appli
34	65.5	8.1	444	3	US-08-464-203-2	Sequence 2, Appli
35	65.5	8.1	444	4	US-08-464-203-2	Sequence 2, Appli
36	65.5	8.1	444	4	US-09-189-607-2	Sequence 2, Appli
37	65.5	8.1	444	4	US-09-378-907-2	Sequence 2, Appli
38	65.5	8.1	444	5	PCT-US94-05779-2	Sequence 2, Appli
39	65.5	8.1	466	1	US-07-882-202A-4	Sequence 4, Appli
40	65.5	8.1	466	1	US-08-021-615A-4	Sequence 4, Appli
41	65.5	8.1	466	1	US-08-321-777-4	Sequence 4, Appli
42	65.5	8.1	466	4	US-09-009-217-14	Sequence 14, Appli
43	65.5	8.1	466	4	US-09-009-656-14	Sequence 14, Appli
44	65.5	8.1	466	5	PCT-US93-04493-4	Sequence 4, Appli
45	65.5	8.1	1241	4	US-09-040-774-2	Sequence 2, Appli

## ALIGNMENTS

RESULT 1  
US-08-563-148E-6  
Sequence 6, Application US/08563148E  
Patent No. 5929224  
GENERAL INFORMATION:  
APPLICANT: Kazuo SUZUKI et al.  
TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESS: 663-2, Shitagi  
STREET: Misaki-machi  
CITY: Isumi-gun  
STATE: Chiba-ken  
COUNTRY: JAPAN  
ZIP: 299-45  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS v.5  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/563 148E  
FILING DATE: No. 5929224ember 27, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 293233/1994  
FILING DATE: 28-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: C. Bruce Hamburg  
REGISTRATION NUMBER: 22,389  
REFERENCE/DOCKET NUMBER: F-5230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)986-2340  
TELEFAX: (212)953-7733  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: single strand  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: intermediate fragment  
ORIGINAL SOURCE:  
ORGANISM: human  
TISSUE TYPE: T-cell derived leukemia cells  
PUBLICATION INFORMATION:  
AUTHORS: Kazuo SUZUKI et al.  
TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 to 151  
US-08-563-148E-6

Query Match 99.9%; Score 806; DB 2; Length 151;  
Best Local Similarity 99.3%; Pred. No. 3.1e-88;  
Matches 150; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGGQYSAORSQRPBGVDILC 60  
DB 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGGQYSAORSQRPBGVDILC 60  
QY 61 SAGSTVYAPFTGMIVGOEKPYQKNAINNGVRISGRGFCVMFYIKPIYKGPICKGEKL 120  
DB 61 SAGSTVYAPFTGMIVGOEKPYQKNAINNGVRISGRGFCVMFYIKPIYKGPICKGEKL 120  
QY 121 GTLLPLQKVPYPGIOSHVHIENCSSDPTAYL 151  
DB 121 GTLLPLQKVPYPGIOSHVHIENCSSDPTAYL 151

RESULT 2  
US-09-194-139-1  
Sequence 1, Application US/09194139A  
Patent No. 6306608  
GENERAL INFORMATION:  
APPLICANT: Arai, Takao  
TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,  
TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME  
FILE REFERENCE: US application 09/194,139  
CURRENT APPLICATION NUMBER: US/09/194,139A  
CURRENT FILING DATE: 1998-11-25  
EARLIER APPLICATION NUMBER: JP 8-132160  
EARLIER FILING DATE: 1996-05-27  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: residue 58 is either Valine or Isoleucine  
US-09-194-139-1

Query Match 99.4%; Score 802; DB 4; Length 151;  
Best Local Similarity 99.3%; Pred. No. 9.4e-88;  
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGGQYSAORSQRPBGVDILC 60  
DB 1 MFSTKALLAGLISALAGPWANICAGKSSNEIRTCDRHGGGQYSAORSQRPBGVDILC 60  
QY 61 SAGSTVYAPFTGMIVGOEKPYQKNAINNGVRISGRGFCVMFYIKPIYKGPICKGEKL 120  
DB 61 SAGSTVYAPFTGMIVGOEKPYQKNAINNGVRISGRGFCVMFYIKPIYKGPICKGEKL 120  
QY 121 GTLLPLQKVPYPGIOSHVHIENCSSDPTAYL 151  
DB 121 GTLLPLQKVPYPGIOSHVHIENCSSDPTAYL 151

RESULT 3  
US-07-924-753-15  
Sequence 15, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESS: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Displaywrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEX:  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-15

Query Match 79.9%; Score 645; DB 1; Length 133;  
Best Local Similarity 87.2%; Pred. No. 3.8e-69;  
Matches 116; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 19 GPWANCAGKSSNEIRTCDRHGGGQYSAORSQRPBGVDILCSAGSTVYAPFTGMIVGOE 78  
DB 1 GPWANCAGKSSNEIRTCDRHGGGQYTAQRNOKLHGVDILCSDSSTVYAPFTGMIVGOE 60



TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: N-terminal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-1

Query Match 24.3%; Score 196; DB 1; Length 41;  
Best Local Similarity 82.9%; Pred. No. 2.1e-16;  
Matches 34; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
OY 19 GPMANICAGKSSNEIRTCDRHGGCGOYSASORSRPHQGVDL 59  
DB 1 GPMATICAGKSSNEIRTCDRHGGCGOYTAAQNRNKLHGQVDVL 41

RESULT 7  
US-07-924-753-14  
Sequence 14, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wandering, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: Displaywrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-14

Query Match 16.4%; Score 132; DB 1; Length 27;  
Best Local Similarity 92.6%; Pred. No. 5e-09;  
Matches 25; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
OY 111 KPIRKGEKLTLLPLQKVPYGIQSHV 137  
DB 1 KGIKKGEKLTLLPLQKVPYGIQSHI 27

RESULT 8  
US-08-563-148E-5  
Sequence 5, Application US/08563148E  
Patent No. 5929224  
GENERAL INFORMATION:  
APPLICANT: Kazuo SUZUKI et al.



;; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
;; NUMBER OF SEQUENCES: 12  
;; CORRESPONDENCE ADDRESSES: 12  
;; ADDRESSEE: 663-2, Shiqi  
;; STREET: Misaki-machi  
;; CITY: Isumi-gun  
;; STATE: Chiba-ken  
;; COUNTRY: JAPAN  
;; ZIP: 299-45  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: MS-DOS v.5  
;; SOFTWARE: Word Perfect 5.1  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/563,148E  
;; FILING DATE: NO. 5929224ember 27, 1995  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: JP 293233/1994  
;; FILING DATE: 28-NOV-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: C. Bruce Hamburg  
;; REGISTRATION NUMBER: 22,389  
;; REFERENCE/DOCKET NUMBER: F-5230  
;; TELEPHONE: (212)986-2340  
;; TELEFAX: (212)953-7733  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single strand  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: intermediate fragment  
;; ORIGINAL SOURCE:  
;; ORGANISM: human  
;; TISSUE TYPE: T-cell derived leukemia cells  
;; PUBLICATION INFORMATION:  
;; AUTHORS: Kazuo SUZUKI et al.  
;; TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
;; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 to 20  
;; US-08-563-148E-5  
Query Match 11.8%; Score 95; DB 2; Length 20;  
Best Local Similarity 85.0%; Pred. No. 8.5e-05;  
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 129 VYPGIQSHVHENCDSPT 148  
DB 1 VYPGIQSHIHENXDLSDPT 20  
RESULT 9  
US-07-924-753-4  
;; Sequence 4, Application US/07924753  
;; Patent No. 5270303  
;; GENERAL INFORMATION:  
;; APPLICANT: Fujio SUZUKI et al.  
;; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
;; NUMBER OF SEQUENCES: 15  
;; CORRESPONDENCE ADDRESSES:  
;; ADDRESSEE: Wenderoth, Lind & Ponack  
;; STREET: 805 Fifteenth Street, N.W., #700  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: U.S.A.  
;; ZIP: 20005  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette, 5.25 inch, 500 KB  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: MS-DOS  
;; SOFTWARE: DisplayWrite

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/924,753  
;; FILING DATE: 19920804  
;; CLASSIFICATION: 424  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Warren M. Cheek, Jr.  
;; REGISTRATION NUMBER: 33,367  
;; REFERENCE/DOCKET NUMBER:  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 202-371-8850  
;; TELEFAX:  
;; INFORMATION FOR SEQ ID NO: 4:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 18 amino acids  
;; TYPE: AMINO ACID  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL:  
;; ANTI-SENSE:  
;; FRAGMENT TYPE: internal fragment  
;; ORIGINAL SOURCE:  
;; ORGANISM: bovine  
;; STRAIN:  
;; INDIVIDUAL ISOLATE:  
;; DEVELOPMENTAL STAGE:  
;; HAPLOTYPE:  
;; TISSUE TYPE: fetal cartilage  
;; CELL TYPE:  
;; CELL LINE:  
;; ORGANELLE:  
;; IMMEDIATE SOURCE:  
;; LIBRARY:  
;; CLONE:  
;; POSITION IN GENOME:  
;; CHROMOSOME/SEGMENT:  
;; MAP POSITION:  
;; UNITS:  
;; FEATURE:  
;; NAME/KEY:  
;; LOCATION:  
;; IDENTIFICATION METHOD:  
;; OTHER INFORMATION:  
;; PUBLICATION INFORMATION:  
;; AUTHORS:  
;; TITLE:  
;; JOURNAL:  
;; VOLUME:  
;; ISSUE:  
;; PAGES:  
;; DATE:  
;; DOCUMENT NUMBER:  
;; FILING DATE:  
;; PUBLICATION DATE:  
;; RELEVANT RESIDUES IN SEQ ID NO: \*  
;; US-07-924-753-4  
Query Match 11.6%; Score 94; DB 1; Length 18;  
Best Local Similarity 88.9%; Pred. No. 9.7e-05;  
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 129 VYPGIQSHVHENCDSPT 146  
DB 1 VYPGIQSHIHENXDLSD 18  
RESULT 10  
US-07-924-753-3  
;; Sequence 3, Application US/07924753

Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEetical:  
ANTI-SENSE:  
FRAGMENT TYPE: Internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:

FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-3  
Query Match 11.2%; Score 90; DB 1; Length 19;  
Best Local Similarity 88.9%; Pred. No. 0.00031;  
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 53 HOGVDLCSAGSTVYAPF 70  
DB 2 HOGVDLCSAGSTVYAPF 19  
RESULT 11  
US-07-924-753-12  
Sequence 12, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEtical:  
ANTI-SENSE:  
FRAGMENT TYPE: Internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:

CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-12

Query Match 11.0%; Score 89; DB 1; Length 19;  
Best Local Similarity 84.2%; Pred. No. 0.00041;  
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 68 APTGMTVGEKPYONKNA 86  
|||||:|||||:|||||  
Db 1 APTGKIMGGEKPYKNKNA 19

RESULT 12  
US-07-924-753-11  
Sequence 11, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: Single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-11

Query Match 10.7%; Score 86; DB 1; Length 18;  
Best Local Similarity 94.1%; Pred. No. 0.00087;  
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 79 KPYONKNAINNGVRISG 95  
|||||:|||||:|||||  
Db 1 KPYKNKNAIINNGVRISG 17

RESULT 13  
5240838-15  
Patent No. 5240838  
APPLICANT: LEDBOER, ADRIANUS M.; MAAT, JAN; VERRIPS, CORNELIS P.  
T.: VISSER, CHRISTIAAN; JANOWICZ, ZBIGNIEW A.; HOLLEMBERG, CORNELIS P.  
TITLE OF INVENTION: REGULATORY SEQUENCES OF ALCOHOL OXIDASE  
(MOX) AND DIHYDROXYACETONESYNTASE (DAS) OF HANSENULA POLYMORPHA  
NUMBER OF SEQUENCES: 19  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/587,555  
FILING DATE: 24-SEP-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 300,211  
FILING DATE: 23-JAN-1989  
APPLICATION NUMBER: 759,315  
FILING DATE: 26-JUL-1985  
SEQ ID NO: 15  
LENGTH: 702  
5240838-15

Query Match 9.2%; Score 74; DB 6; Length 702;  
Best Local Similarity 24.1%; Pred. No. 3.1;  
Matches 33; Conservative 22; Mismatches 48; Indels 34; Gaps 7;

QY 11 GLISTALAGPWANICAGKSSNEIRTCDRHCGGYSAGRSORPHOGVDILCSAGSTVAPF 70

Db 525 GLFOKAVELPSSILS-LSRNEVL-----OYLASRQRRRNAGYILLEDAENAEVQI 575  
Oy 71 TGMIVGQEKRYONKNAINNGVIRISGRF-----CYKMFYIKPKYKGP1--KKGKEL 120  
Db 576 IG--VGAEMFEADKA-----KILGRKFRTRVLSIPCRLLFDEOSIGYRSRYLRKGRQV 628  
Oy 121 GTLLPLQKVPYGIOSHV 137  
Db 629 PTIV-----VDARV 637

RESULT 14  
US-07-924-753-10  
Sequence 10, Application US/07924753  
Patent No. 5270303

GENERAL INFORMATION:  
APPLICANT: FUJIO SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:

TELEX:  
INFORMATION FOR SEQ. ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: Internal fragment  
ORIGINAL SOURCE:  
ORGANISM: Bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:

UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

US-07-924-753-10

Query Match 9.0%; Score 73; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 119 KGTLPLQKVPYGI 132  
Db 1 KGTLPLQKVPYGI 14

RESULT 15  
US-08-996-139-15  
Sequence 15, Application US/08996139  
Patent No. 6017729

GENERAL INFORMATION:  
APPLICANT: Anderson, Dirk M.  
APPLICANT: Maraskovsky, Eugene  
TITLE OF INVENTION: Receptor Activator of NF-kappaB  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation, Law Department  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Power Macintosh 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/996,139  
FILING DATE: 22 DECEMBER 1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 60/064,671  
FILING DATE: 14 OCTOBER 1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/813,509  
FILING DATE: 07 MARCH 1997  
APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/772,330  
FILING DATE: 23 DECEMBER 1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,693  
REFERENCE/DOCKET NUMBER: 2851-A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)233-0644  
INFORMATION FOR SEQ. ID NO: 15:  
SEQUENCE CHARACTERISTICS:

; LENGTH: 625 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-996-139-15

Query Match 8.7%; Score 70.5; DB 3; Length 625;  
Best Local Similarity 21.4%; Pred. No. 7;  
Matches 31; Conservative 12; Mismatches 39; Indels 63; Gaps 5;

QY 16 ALAGPWANI-----CAGKSNIEIRTCDRHRCGCGQYSAQRSGORPHQGVDIICSA 62  
DB 309 AAGGPWAEVRDSCRFTLVSEVETQGLSKRIPTEDY-----TDRPSQSTGSLITQO 362  
QY 63 GSTVYAPFTGMIVGQEKPYQNNAINNGVRIISGRGFCVKMFYIKPIKKGPPIKKGEKLG 122  
DB 363 GSKSIPPF-----QE-----PLEVGENDS- 381  
QY 123 LLPLQKYVPGIQSHVHIENCSSDP 147  
DB 382 ---LSOCFTGTSTVDSEGCDFTEP 403

Search completed: March 25, 2003, 17:23:00  
Job time : 19 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:21:15 ; Search time 12 Seconds  
(without alignments)  
672.733 Million cell updates/sec

Title: US-09-868-953a-2

Perfect score: 807  
Sequence: 1 MESTKALLAGLISTALACP.....GIOSHVHENCSSDPAYL 151

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PCN\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/PCNUS\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	544	67.4	113 10 US-09-864-570-9	Sequence 9, Appli
2	544	67.4	113 12 US-10-157-457-9	Sequence 9, Appli
3	430	53.3	326 10 US-09-864-570-8	Sequence 8, Appli
4	430	53.3	326 12 US-10-157-457-8	Sequence 8, Appli
5	352	43.6	115 10 US-09-864-570-11	Sequence 11, Appli
6	352	43.6	115 12 US-10-157-457-11	Sequence 11, Appli
7	347	43.0	115 10 US-09-864-570-10	Sequence 10, Appli
8	347	43.0	115 12 US-10-157-457-10	Sequence 10, Appli
9	294	36.4	54 10 US-09-864-761-38651	Sequence 10, Appli
10	82	10.2	25 10 US-09-864-570-7	Sequence 38651, A
11	82	10.2	25 12 US-10-157-457-7	Sequence 7, Appli
12	74	9.2	22 10 US-09-864-570-5	Sequence 5, Appli
13	74	9.2	22 12 US-10-157-457-5	Sequence 5, Appli
14	73	9.0	16 10 US-09-864-570-6	Sequence 6, Appli
15	73	9.0	16 12 US-10-157-457-6	Sequence 6, Appli
16	72.5	8.7	168 10 US-09-738-626-5722	Sequence 5722, Ap
17	70.5	8.7	625 9 US-09-877-650-15	Sequence 15, Appli
18	70.5	8.7	625 9 US-10-166-232A-2	Sequence 15, Appli
19	70.5	8.7	625 10 US-09-871-856-15	Sequence 15, Appli

20	67.5	8.4	2285 10 US-09-932-183A-2	Sequence 2, Appli
21	65.5	8.1	406 9 US-10-109-498-1	Sequence 1, Appli
22	65.5	8.1	502 9 US-09-738-626-4722	Sequence 4722, Ap
23	65.5	8.1	1241 10 US-09-759-622-2	Sequence 2, Appli
24	65	8.1	412 10 US-09-815-242-5451	Sequence 5451, Ap
25	65	8.1	412 10 US-09-815-242-12381	Sequence 12381, A
26	64.5	8.0	1080 10 US-09-904-380-2	Sequence 2, Appli
27	63.5	7.9	151 10 US-09-925-301-874	Sequence 874, App
28	63.5	7.9	723 9 US-10-133-912-18	Sequence 18, Appli
29	63	7.8	142 10 US-09-815-242-5718	Sequence 5718, Ap
30	63	7.8	280 10 US-09-815-242-12479	Sequence 12479, A
31	62.5	7.7	332 10 US-09-925-300-1064	Sequence 1064, Ap
32	62.5	7.7	197 9 US-09-924-340-66	Sequence 66, Appli
33	62.5	7.7	197 9 US-09-992-600A-66	Sequence 66, Appli
34	62.5	7.7	197 10 US-09-731-872-276	Sequence 276, App
35	62.5	7.7	375 10 US-09-864-761-36336	Sequence 36336, A
36	62.5	7.7	492 9 US-10-087-667-6	Sequence 6, Appli
37	62.5	7.7	615 10 US-09-768-779A-2	Sequence 2, Appli
38	62.5	7.7	1038 10 US-09-908-500A-2	Sequence 2, Appli
39	62	7.7	269 10 US-09-852-100A-2	Sequence 2, Appli
40	62	7.7	269 10 US-09-833-503A-2	Sequence 2, Appli
41	62	7.7	598 10 US-09-854-845-41	Sequence 41, Appli
42	62	7.7	697 10 US-09-854-845-25	Sequence 25, Appli
43	61.5	7.6	480 9 US-09-895-913A-92	Sequence 92, Appli
44	61.5	7.6	480 9 US-10-087-667-5	Sequence 5, Appli
45	61.5	7.6	480 9 US-10-087-667-8	Sequence 8, Appli

#### ALIGNMENTS

RESULT 1  
US-09-864-570-9  
Sequence 9, Application US/09884570  
Patent No. US20020114779A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294  
CURRENT APPLICATION NUMBER: US/09/884,570  
PRIOR FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/212,271  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 9  
LENGTH: 113  
TYPE: PRT  
ORGANISM: unknown  
FEATURE:  
LOCATION: 38..165  
OTHER INFORMATION: amino acid sequence of region of lect2  
OTHER INFORMATION: protein homologous to repeat sequences of N  
US-09-864-570-9

Query Match 67.4%; Score 544; DB 10; Length 113;  
Best Local Similarity 85.8%; Pred. No. 6.3e-54;  
Matches 97; Conservative 10; Mismatches 6; Indels 0; Gaps 0;  
QY 39 HGGCGYSAGRSORPHOGVDILCSAGSTVYAFPTGMIVGOEKRYONKAINNGVRSIGRGF 98  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 YGGCGYSAGRTGRNHNHGVYVLCSDGSVYVAFPTGKIYGOEKRYNRKKNAINDSIRLSGRGF 60  
QY 99 CVKMYIPIKTKKPKKKEKLTLLPLQKVYVPGISQSHVHENCSSDPAYL 151  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 CVKIFYIPIKYKGSIKKKEKLTLLPLQKTYVIOQSHVHENCSSDPAYL 113

RESULT 2  
US-10-157-457-9  
Sequence 9, Application US/10157457







```
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38651
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004763.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8
; OTHER INFORMATION: SWISSPROT HIT: O14960, EVALU8 8.00e-17
US-09-864-761-38651
```

```
Query Match 36.4%; Score 294; DB 10; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.4e-26;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 98 FCVKMEYIKPIKYGKIKGKIGTLPLQKVPYGIQSHVHIENCDSDPTAYL 151
Db 1 FCVKMEYIKPIKYGKIKGKIGTLPLQKVPYGIQSHVHIENCDSDPTAYL 54
```

## RESULT 10

```
US-09-884-570-7
; Sequence 7, Application US/09884570
; Patent No. US20020114779A1
; GENERAL INFORMATION:
; APPLICANT: Williams, John P.
; APPLICANT: McDonald, Jay M.
; APPLICANT: McKenna, Margaret A.
; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294
; CURRENT APPLICATION NUMBER: US/09/884,570
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/212,271
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 11
```

```
; SEQ ID NO 7
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:
; LOCATION: 219..243
; OTHER INFORMATION: amino acid sequence tryptic peptide from
; OTHER INFORMATION: mim-1 protein
US-09-884-570-7
```

```
Query Match 10.2%; Score 82; DB 10; Length 25;
Best Local Similarity 56.0%; Pred. No. 0.0093;
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 55 GVDILCSAGSTYVAPFTGMIVGOEK 79
Db 1 GVDIVCAGATVYAPFSGELSGPVK 25
```

```
RESULT 11
US-10-157-457-7
; Sequence 7, Application US/10157457
; Patent No. US20020150568A1
; GENERAL INFORMATION:
; APPLICANT: Williams, John P.
; APPLICANT: McDonald, Jay M.
; APPLICANT: McKenna, Margaret A.
; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294CIP
; CURRENT APPLICATION NUMBER: US/10/157,457
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: US 09/884,570
; PRIOR FILING DATE: 2001-06-19
; NUMBER OF SEQ ID NOS: 11
; SEQ ID NO 7
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:
; LOCATION: 219..243
; OTHER INFORMATION: amino acid sequence tryptic peptide from
; OTHER INFORMATION: mim-1 protein
US-10-157-457-7
```

```
Query Match 10.2%; Score 82; DB 12; Length 25;
Best Local Similarity 56.0%; Pred. No. 0.0093;
Matches 14; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 55 GVDILCSAGSTYVAPFTGMIVGOEK 79
Db 1 GVDIVCAGATVYAPFSGELSGPVK 25
```

## RESULT 12

```
US-09-884-570-5
; Sequence 5, Application US/09884570
; Patent No. US20020114779A1
; GENERAL INFORMATION:
; APPLICANT: Williams, John P.
; APPLICANT: McDonald, Jay M.
; APPLICANT: McKenna, Margaret A.
; TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof
; FILE REFERENCE: D6294
; CURRENT APPLICATION NUMBER: US/09/884,570
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/212,271
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 11
; SEQ ID NO 5
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Gallus gallus
; FEATURE:
```

LOCATION: 94.115  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-09-884-570-5

Query Match  
Best Local Similarity 61.9%; Pred. No. 0.063;  
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 81 YONKNAINGVRISGRFCVK 101  
DB 2 FHNGNAIDGVQISGSYCVK 22

RESULT 13  
US-10-157-457-5  
Sequence 5, Application US/10157457  
Patent No. US20020150568A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294CIP  
CURRENT APPLICATION NUMBER: US/10/157,457  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: US 09/884,570  
PRIOR FILING DATE: 2001-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 5  
LENGTH: 22  
TYPE: PRT  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 94.115  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-10-157-457-5

Query Match  
Best Local Similarity 9.2%; Score 74; DB 12; Length 22;  
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 81 YONKNAINGVRISGRFCVK 101  
DB 2 FHNGNAIDGVQISGSYCVK 22

RESULT 14  
US-09-884-570-6  
Sequence 6, Application US/09884570  
Patent No. US20020114779A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294  
CURRENT APPLICATION NUMBER: US/09/884,570  
CURRENT FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/212,271  
PRIOR FILING DATE: 2000-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 6  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 293.308  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-09-884-570-6

Query Match  
Best Local Similarity 9.0%; Score 73; DB 10; Length 16;  
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 129 VYPGIOSHVHENC D 143  
DB 1 VFPGLISHVHENC D 15

RESULT 15  
US-10-157-457-6  
Sequence 6, Application US/10157457  
Patent No. US20020150568A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294CIP  
CURRENT APPLICATION NUMBER: US/10/157,457  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: US 09/884,570  
PRIOR FILING DATE: 2001-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 6  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 293.308  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-10-157-457-6

Query Match  
Best Local Similarity 9.0%; Score 73; DB 12; Length 16;  
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 129 VYPGIOSHVHENC D 143  
DB 1 VFPGLISHVHENC D 15

Search completed: March 25, 2003, 17:29:45  
Job time : 13 secs



GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:16:59 ; Search time 28.5 Seconds  
(without alignments)  
509.344 Million cell updates/sec

Title: US-09-868-953A-2

Perfect score: 807  
Sequence: 1 MFSTKALLLAGLISTALAGP.....GIQSHVHIENCDSDPTAYL 151

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73: \*  
1: p1r1: \*  
2: p1r2: \*  
3: p1r3: \*  
4: p1r4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	645	79.9	133	2	chondromodulin II
2	433	53.7	326	2	myb-induced myeloid
3	103.5	12.8	305	2	hypothetical prote
4	80.5	10.0	417	2	conserved hypotnet
5	79	9.8	750	2	lipoprotein lipor
6	77.5	9.6	475	2	hypothetical prote
7	76	9.4	651	2	hypothetical prote
8	75	9.3	240	2	hypothetical prote
9	75	9.3	444	2	hypothetical prote
10	74.5	9.2	268	2	probable membrane
11	74.5	9.2	295	2	hypothetical prote
12	74	9.2	284	2	probable periplasm
13	73	9.0	350	2	hypothetical prote
14	73	9.0	710	1	formaldehyde trans
15	73	9.0	1331	1	xanthine dehydroge
16	72.5	9.0	355	2	CONSTANS-like 1
17	72	8.9	1434	2	hypothetical prote
18	70.5	8.7	428	2	probable penicilli
19	70.5	8.7	454	2	formyl-coenzyme a
20	70.5	8.7	706	2	subtilisin-like pr
21	70	8.7	689	2	hypothetical prote
22	69	8.6	310	2	insulin-like growt
23	69	8.6	1347	2	helicase homolog F
24	68.5	8.5	160	2	hypothetical prote
25	68.5	8.5	233	2	hypothetical prote
26	68.5	8.5	273	2	probable periplasm
27	68.5	8.5	357	2	isopropylmalate de
28	68.5	8.5	608	2	gtp-binding protei
29	68	8.4	184	2	hypothetical prote

30	68	8.4	407	2	probable pectinest
31	68	8.4	762	2	DNA mismatch repai
32	67.5	8.4	312	2	hypothetical prote
33	67.5	8.4	443	2	coagulation factor
34	67.5	8.4	535	2	hypothetical prote
35	67.5	8.4	535	2	Rf1 protein - saim
36	67.5	8.4	2285	2	conserved transglyc
37	67	8.3	271	2	conserved hypothet
38	67	8.3	609	2	peptidase, M23/M37
39	66.5	8.2	1012	2	DNA polymerase - h
40	66	8.2	271	2	probable lipoprote
41	66	8.2	335	2	probable transposa
42	66	8.2	398	2	integumentary muc1
43	66	8.2	426	2	hypothetical prote
44	66	8.2	447	2	conserved hypothet
45	66	8.2	554	2	component of type

#### ALIGNMENTS

##### RESULT 1

JH0270 Chondromodulin II - bovine

N/Alternate names: cartilage-derived factor

C/Species: Bos primigenius taurus (cattle)

C/Date: 07-Jul-1997 #sequence\_revision 18-Jul-1997 #text\_change 17-Mar-1999

R/Hiraki, Y.; Inoue, H.; Kondo, J.; Kamizono, A.; Yoshitake, Y.; Shukunami, C.;

J. Biol. Chem. 271, 22657-22662, 1996

A/Title: A novel growth-promoting factor derived from fetal bovine cartilage, c

A/Reference number: JH0270; MUID:96394331; PMID:8798437

A/Accession: JH0270

A/Molecule type: protein

A/Residues: 1-133 <HR>

A/Experimental source: cartilage cell

C/Comment: This protein stimulates proteoglycan synthesis in the growth plate c

Query Match 79.9%: Score 645; DB 2: Length 133;

Best local similarity 87.2%; Pred. No. 7.3e-58;

Matches 116; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY	19	GPWANICAGKSSNEIRTCDRHGGQYSAORSORPHQGVDLCSAGSTVYAPFTGMVGOE	78
DB	1	GPWANICAGKSSNEIRTCDRHGGQYTAQNRKLGQVDVLCSDGSTVYAPFTGKIMGGE	60
QY	79	KPYQKNAINNGVRISGRFCYKMFYIKPKIKYGPRIKKEKLGTLPLQKYYPGIIQSHVH	138
DB	61	KPYKKNAINNGVRISGGGFCIKMFYIKPKIKYKSIKKEKLGTLPLQKYYPGIIQSHIH	120
QY	139	IENCDSPTAYL 151	
DB	121	IENCDSPTAYL 133	

##### RESULT 2

A33755 myb-induced myeloid protein 1 (mim-1) precursor - chicken

N/Alternate names: arginine-specific ADP-ribosyltransferase endogenous substrat

C/Species: Gallus gallus (chicken)

C/Date: 30-Mar-1990 #sequence\_revision 30-Mar-1990 #text\_change 04-Sep-1998

C/Accession: A33755; S29229

R/Ness, S.A.; Marknell, A.; Graf, T.

A/Title: The v-myb oncogene product binds to and activates the promyelocyte-spe

A/Reference number: A33755; MUID:90090611; PMID:2688896

A/Accession: A33755

A/Molecule type: mRNA

A/Residues: 1-326 <NES>

A/Cross-references: GB:M29449; MUD:9212341; PID:9212342

R/Tamada, K.; Tsuchiya, M.; Mishima, K.; Shimoyama, M.

FEBS Lett. 311, 203-205, 1992

A/Title: p13, an endogenous target protein for arginine-specific ADP-ribosyltran

[illegible]

A:Accession: C82840  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-417 <SIM>  
A:Cross-references: GB:AE003870; GB:AE003849; NID:g9104955; PIDN:AAF82980.1; GSP:  
R:Experimental source: strain 965C  
R:Simpon, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvareng  
B:Rios, M.R.S.; Bueno, M.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; C  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kurame, E.E  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Ma  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmit  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.;  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da  
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
C:Superfamily: hypothetical protein H10409  
Query Match 10.0%; Score 80.5; DB 2; Length 417;  
Best Local Similarity 25.0%; Pred. No. 2.5;  
Matches 29; Conservative 24; Mismatches 42; Indels 21; Gaps 7;  
OY 34 PTCDDHGGCGYAAORSQRPDHOGVDILCSAGSYTAAFTGMI--VGOEKPYONKNAINGV 91  
DB 249 RLSSTFFGLRRHPIILGMRHKKGVDAATAGTPIMAAGADARVGTIGDQGRGNYIILH-- 306  
OY 92 RISGRGFCYKMYIPIPIKYKGRPKKGEKL--GTLPLQKYVPIQS----HWITE 140  
DB 307 ---GKY--RTLIGHMSRF--GRIKAGOKINOGTVIG---YVGMTGLATPRLHYE 352  
RESULT 5  
F90571  
I:lipoprotein [imported] - Mycoplasma pulmonis (strain UAB C11P)  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #ext-change 03-Aug-2001  
R:Chambaud, I.; Hellio, R.; Ferris, S.; Barbe, V.; Samson, D.; Gallison, F.; Mos  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplas  
A:Reference number: A59512; MUID:21267165; PMID:11353084  
A:Accession: F90571  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-750 <KUB>  
A:Cross-references: GB:AL445566; PID:g14089892; PIDN:CA03651.1; GSPDB:GN00153  
A:Experimental source: strain UAB C11P  
C:Genetics:  
A:Gene: MYPu\_4780  
A:Genetic code: SGC3  
Query Match 9.8%; Score 79; DB 2; Length 750;  
Best Local Similarity 22.8%; Pred. No. 6.4;  
Matches 42; Conservative 25; Mismatches 55; Indels 62; Gaps 8;  
OY 12 LISTALAGWNAICAGKSNSNIRTCDR--HGCGYAAORSQRPDHOGVDILCSAGSYTAA 69  
DB 456 LQSPVSGWMD-----DRTSYSTSKADSKNTKDIHLGEDILILEONKEVIAP 502  
OY 70 FTGMTVGQ--EKPYO-----KNNAINGVRIISGRGFCYKFF-- 103  
DB 503 FDGKIIASYYAASPOAYGLGVITVLEWKKDLVGOIDSVINDIOLAEFTRIYIAFHLN 562  
OY 104 -----YIKPIYKKG-----PIKGEKIGTLLPLQKYVAGIOSHVIIE--NCD 143  
DB 563 PSYLENYGLVLEVOSSTAIEITPATPKYVKGCEVIG--LVGEKRNNGGMKPHVHIEVSLG 621



RESULT 10  
D97054  
probable membrane metalloendopeptidases [imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
A:Accession: D97054  
R:Noiling, J.; Breton, G.; Omelchenko, M.V.; Makarova, K.S.; Zeng, Q.; Gibson, R.; Lee,  
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
J. Bacteriol. 183, 4823-4838, 2001  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cld  
A:Reference number: A96900; MVID:21359325; PMID:21359325  
A:Accession: D97054  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-268 <KUR>  
A:Cross-references: GB:AE001437; PTDN:AAK79223.1; PTD:q15024178; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC824  
C:Genetics:  
A:Gene: CAC1252

Query Match	Similarity	9.2%	Score 74.5	DB 2	Length 268
Best Local	Similarity	23.8%	Pred. No. 6.5		
Matches	31	Conservative	29	Mismatches	51
				Indels	19
				Gaps	6
QY	27	GKSSNEIRTCDRHGGCGVYSAQRSGRPHOGVDILCSAGSYVAPFTGNI--VGQEKPYONK	84		
		: : : : :   : : : : :   : : : : :   : : : : :   : : : : :			
DB	148	GKLSAHPKQKQDKDVLG-----NQYLNGVDIKANSDDLVAIKALYDGTVEKVGEDKDPGY	201		
QY	85	NAINNGVRIISGRGCVKMFYIKPIYKGP-IKKGEKLGTLPLPÖKYYPGİQSHVHIENC-	142		
		: : : : :   : : : : :   : : : : :   : : : : :   : : : : :			
DB	202	VAMDNGGVGVESKYSNMDSFEVQ---RGDGVYKGEVLG--KYKKINDAKASKSYLHFEIMY	254		
QY	143	--DSDPTAY	150		
		: : : : :			
DB	255	MGENDPENT	264		

```

RESULT 11
AC2451
hypothetical protein al15163 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp., Strain PCC 7120 is a synonym of Anabaena sp., strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AC2451
R:Kaneiko, T.; Nakamura, Y.; Wolz, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Iriuchihara, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AC2451
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-295 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAF76862.1; PID:gl7134301; GSPDB:GN00179
A:Experimental source: Strain PCC 7120
C:Genetics:
A:Gene: al15163

```

```

Query Match Similarity      9.28; Score 74.5; DB 2; Length 295;
Best Local Similarity      29.68; Pred. No. 7.1;
Matches      34; Conservative      15; Mismatches      39; Indels      27; Gaps      7;

QY      53 HGGVDILCSAGSTV--VAPFTGMIVGOEKPYPONKNAIN-NGVRISGRGFCVKMYEYIKR      109
      | | | | | : : : | | | | : : | | | | : : |
Db      190 HSCVDLLAALGCTPYEAIAIGTVFKAKEQTYNLVYIINHG-----GLQSRVAHLDSIN      243

QY      110 YK--GPIKKEGLTGL---LPLQKVPYEGIQSHVHIENQDS-----DPTAYL      151
      : : : : | : | : | : | : | : | : | : | : |
Db      244 VKYGGKVMQCGLLGTGTTGQPTAK-----QPHLEFVRYSSSLGVAENPKDYL      293

```

RESULT 12  
E97836

Probable peptidylase protein [imported] - Rickettsia conorii (strain Malish 7)  
 C:Species: Rickettsia conorii  
 C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001  
 C:Accession: E97836  
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fourmiller, P.E.; Barde, V.; Samson  
 Science 293, 2093-2098, 2001  
 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii  
 A:Reference number: A97700; MUID:21442074; PMID:11557893  
 A:Accession: E97836  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-284 <KUR>  
 A:Cross-references: GB:AE006914; PIDN:AAL03631.1; PID:g15620216; GSPDB:GN00173  
 C:Genetics:  
 A:Gene: RCI093

	Query Match	Score	DB	Length
Best Local Similarity	24.0%	74	2	284
Matches	24	Conservative	17	Mismatches
			31	Indels
			28	Gaps
			5	

```

Oy  53 HGGVDFLLCSGSGVYAPFAFTGIV--GGGRPKPNKMAINNGVIRISGCEFCVMKVIPIKY 110
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  181 HSGIDLDQAKKAPDYIAAASGIVIKAAAPDYGNFEVEIKH-----GRRFTYKVAHLKEMSV 235
Oy  111 K--GPIKKGKELGTLPLQKVPYRQ-----SHVHE 140
    1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  236 KEGNKIKTRGQ-----FTIDQSGTGNAGTGEHLHE 264

```

RESULT 13  
AD0466  
hypothetical protein YPO3828 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
C:Accession: AD0466  
R:Parkhill, J., Wren, B.W., Thomson, N.R., Tittball, R.W., Holden, M.T.G., Prend  
leno-Farrago, A.M., Chillingworth, T., Cronin, A., Davies, R.M., Davis, P., Do  
ll, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S.,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AD0466  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-350 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC93296.1; PID:g15981743; GSPDB:GM00175  
A:Genetics:  
.:Gene: YPO3828

Query Match	Similarity	9.0%	Score 73:	DB 2:	length 350:
Best Local	Similarity	26.3%	Pred. No. 12:		
Matches	25:	Conservative	17:	Mismatches	39:
				Indels	14:
				Gaps	4
QY	52	PHQGVDFILCSASSTYVAPE-----TCGMYIGOEKRPYKNAINNGVIRISGRGFCVKKFYIK	106		
DB	210	PP-----CHGFDELATASIKVTDKOSCHITFEELKLYSTMANISMTIIFPEKGYIKLIDIFHAE	263		
QY	107	PIKYKGPPIKGEKLTLLPLQKVPYEGISQHVHLEN	141		
DB	264	PSRLR--VSSGDS-DTLIPESKONSTIMTISIGLEN	295		

RESULT 14  
KJ0HFK  
Formaldehyde transketolase (EC 2.2.1.3) - yeast (*Pichia angusta*)  
N:Alternate names: dihydroxyacetone synthase; glycerone synthase  
C:Species: *Pichia angusta*  
C:Date: 31-Mar-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Apr-2000  
C:Accession: A23009; S30110  
R:Janowicz, Z. A.; Eckart, M. R.; Drewke, C.; Roggenskamp, R. O.; Hollenberg, C. P.  
Nucleic Acids Res. 13, 3043-3062, 1985  
A:Title: Cloning and characterization of the DAS gene encoding the major metha  
:Reference number: A23009; MUID:85215670; PMID:2987872



A:Accession: A23009  
 A:Molecule type: DNA  
 A:Title: *Rhipicephalus* 'JAN'  
 A:Cross-references: GB:X02424  
 A:Experimental source: ATCC 34438  
 A>Note: this sequence has been revised in reference S30110  
 R.Hansen, H.; Didion, T.; Thiemann, A.; Veenhuis, M.; Roggenkamp, R.  
 Mol. Gen. Genet. 235, 269-278, 1992  
 A:Title: Targeting sequences of the two major peroxisomal proteins in the methylotrophic  
 A:Reference number: S30110; MUID:93101130; PMID:1465101  
 A:Accession: S30110  
 A:Molecule type: DNA  
 A:Residues: 667-710 <HAN>  
 A:Cross-references: EMBL:X02424  
 A>Note: this is a revision to the sequence from reference A23009  
 C:Comment: This is the major methanol assimilatory enzyme from this methylotrophic organ  
 C:Genetics:  
 A:Gene: DAS  
 C:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology  
 C:Keywords: peroxisome; thiamin pyrophosphate; transferase  
 F:158-208/Domain: thiamin pyrophosphate-binding domain homology #status atypical <TPB>  
 F:708-710/Region: peroxisome/glyoxysome location signal #status atypical

Query Match 9.0%; Score 73; DB 1; Length 710;  
 Best Local Similarity 24.1%; Pred. No. 25;  
 Matches 33; Conservative 21; Mismatches 49; Indels 34; Gaps 7;  
 QY 11 GLISTALAGPWANICAGKSNIEIRTCDRHGGCGYSAQSORPHQGVDTLSAGSTVYAPF 70  
 DB 525 GLFQKAVELPFSSILS-LSRNEVL-----QYLSARQRRNNAAGYILEDENAEYQI 575  
 QY 71 TGMVGOEKPYQNKNAINGVRIISGRF-----CVAMFYIKPIKYKGPV--KKGEKL 120  
 DB 576 IG--VGAEMEFADKAA-----KILGRKFRTRVLSIPCTRLEQSGISGRSVLRKDKDQV 628  
 QY 121 GTLLPLQKVPVGIOSHV 137  
 DB 629 PTVV-----VDGHV 637

## RESULT 15

XORTH  
 xanthine dehydrogenase (EC 1.1.1.204) / xanthine oxidase (EC 1.1.3.22) - rat  
 N:Alternate names: hypoxanthine oxidase  
 C:Species: *Rattus norvegicus* (Norway rat)  
 C:Dates: 30-Apr-1991 #sequence revision 07-Feb-1997 #extl\_change 19-Jan-2001  
 R:Accession: A37810; S45259; S45260; S71397; I58308  
 R:Amaya, Y.; Yamazaki, K.; Sato, M.; Noda, K.; Nishino, T.; Nishino, T.  
 J. Biol. Chem. 265, 14170-14175, 1990  
 A:Title: Proteolytic conversion of xanthine dehydrogenase from the NAD-dependent type to  
 vage sites of the enzyme protein of xanthine dehydrogenase by trypsin.  
 A:Reference number: A37810; MUID:90354396; PMID:2387845  
 A:Accession: A37810  
 A:Molecule type: mRNA  
 A:Residues: 1-478,491-493,'Q',495-1331 <AMA>  
 A:Cross-references: GB:J05579; NID:9207686  
 A>Note: parts of this sequence, including the amino end of the mature protein, were dete  
 R:Chow, C.W.; Clark, M.; Rinaldo, J.; Chalkley, R.  
 Nucleic Acids Res. 22, 1846-1854, 1994  
 A:Title: Identification of the rat xanthine dehydrogenase/oxidase promoter.  
 A:Reference number: I58308; MUID:94268906; PMID:8208609  
 A:Accession: S45259  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 476-494 <RES>  
 A:Cross-references: EMBL:U08123; NID:9473260; PIDN:AA60444.1; PID:9473261  
 A>Note: correction to A37810; sequence thought by authors to be macrophage splice form  
 A:Accession: S45260  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-55 <CHO>  
 A:Cross-references: GB:U08122; NID:9472856; PIDN:AA18869.1; PID:9472858; EMBL:U08121  
 R:Sato, A.; Nishino, T.; Noda, K.; Amaya, Y.; Nishino, T.

J. Biol. Chem. 270, 2818-2826, 1995  
 A:Title: The structure of chicken liver xanthine dehydrogenase. cDNA cloning and  
 A:Reference number: A55711; MUID:95155354; PMID:7852355  
 A:Contents: annotation; confirmation of sequence  
 A>Note: the authors confirmed that both liver and macrophage mRNA's of the rat h  
 R:McManaman, J.L.; Shellman, V.; Wright, R.M.; Repine, J.E.  
 Arch. Biochem. Biophys. 332, 135-141, 1996  
 A:Title: Purification of rat liver xanthine oxidase and xanthine dehydrogenase b  
 A:Reference number: S71397; MUID:96400342; PMID:8806718  
 A:Accession: S71397  
 A:Molecule type: protein  
 A:Residues: 2-11 <MCV>  
 C:Comment: xanthine dehydrogenase is reversibly converted to xanthine oxidase by  
 version to xanthine oxidase can also be performed artificially by a variety of  
 C:Comment: This enzyme contains four cofactors per subunit, one FAD, two iron-su  
 C:Genetics:  
 A:Introns: 14/3; 34/1  
 A>Note: the list of introns may be incomplete  
 C:Complex: homodimer  
 C:Function: <XDH>  
 A:Description: catalyzes oxidation of xanthine to uric acid by NAD+ and water  
 A:Pathway: purine catabolism  
 C:Function: <XO>  
 A:Description: catalyzes oxidation of xanthine to uric acid and hydrogen peroxide  
 A:Pathway: purine catabolism  
 C:Superfamily: xanthine dehydrogenase; ferredoxin [2Fe-2S] homology  
 C:Keywords: 2Fe-2S; FAD; flavoprotein; homodimer; iron-sulfur protein; metallopr  
 11sm  
 F:2-1331/Product: xanthine dehydrogenase / xanthine oxidase #status experimental  
 F:26-74/Domain: ferredoxin [2Fe-2S] homology <FER>  
 F:795-802/Region: nucleotide-binding motif A (P-loop)  
 F:43-48,51,73/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted  
 F:112-115,147,149/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted  
 F:825/Binding site: molybdopterin (Cys) (covalent) #status predicted  
 F:912/Binding site: molybdopterin (Arg) #status predicted  
 F:1261/Active site: Glu #status predicted

Query Match 9.0%; Score 73; DB 1; Length 1331;  
 Best Local Similarity 28.8%; Pred. No. 46;  
 Matches 34; Conservative 12; Mismatches 42; Indels 30; Gaps 6;  
 QY 35 TCDBHGGGQYSA---QSRPHQGVDTLSAGSTVYAPFTGTVGOEKPYQNKNAINGV 91  
 DB 661 TCVGHIIQAVVADPPEHQAARGVKI-----TYEDLPALITIDP-----AINN-- 704  
 QY 92 RISRGCVKMFYIKPIKY-KGPIKKGKGLTLLPLQKVPVGIOSHVHIE-NCDSOP 147  
 DB 705 -----NSFYSGEIKTEKGLKKGFEADNVVSGELTIGQDEHPYLETNCTIAPV 753

Search completed: March 25, 2003, 17:22:10  
 Job time : 29.5 secs



GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:39 ; Search time 8.5 Seconds

(without alignments)  
736.815 Million cell updates/sec

Title: US-09-868-953a-2

Sequence: 1 MFSTKALLAGLISALAGP.....GIQSHVHENCSSDPAYL 151

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	806	99.9	151	1	LCR2_HUMAN
2	702	87.0	151	1	LCR2_BOVIN
3	685	84.9	151	1	LCR2_MOUSE
4	433	53.7	326	1	MTM1_CHICK
5	77.5	9.6	475	1	YEBP_HAEN
6	73	9.0	710	1	DAS_PICAN
7	73	9.0	1330	1	XDH_RAT
8	72.5	9.0	355	1	COL1_ARATH
9	70.5	8.7	625	1	TRIL_MOUSE
10	70.5	8.7	680	1	Z334_HUMAN
11	69.5	8.6	472	1	ZIM3_HUMAN
12	69	8.6	779	1	TRF3_TREVO
13	68	8.4	184	1	YTF63_CAEEL
14	68	8.4	762	1	MUS2_AQUAE
15	67.5	8.4	444	1	PA7_RABIT
16	67.5	8.4	535	1	VS50_HSVSA
17	66.5	8.2	1012	1	DPOL_HSVJ
18	66	8.2	398	1	MOB1_XENLA
19	66	8.2	564	1	HEMA_IADA2
20	65.5	8.1	362	1	FIHP_ADE08
21	65.5	8.1	466	1	PA7_HUMAN
22	65.5	8.1	1241	1	NPHN_HUMAN
23	65.5	8.1	1388	1	REOD_TOBAC
24	65	8.1	317	1	IBP2_SHEEP
25	65	8.1	412	1	GLYA_STAM
26	64.5	8.0	357	1	LEU3_CLOPA
27	64.5	8.0	391	1	HTS2_BACSU
28	64.5	8.0	551	1	LEM2_RABIT
29	64.5	8.0	2241	1	TG0F_HCVNA
30	64	7.9	271	1	YLBG_ECOLI
31	64	7.9	466	1	SELA_RHIME
32	64	7.9	593	1	MTF2_HUMAN
33	64	7.9	746	1	GYP7_YEAST

34	64	7.9	1133	1	VINC_HUMAN	P18206 homo sapien
35	63.5	7.9	571	1	SECD_MYCE	P38387 mycobacteri
36	63.5	7.9	664	1	PTSA_STRMU	P13655 streptococc
37	63	7.8	288	1	TR6_NEIMB	O00840 neisseria m
38	63	7.8	312	1	PGLR_AGRTU	P27644 agrobacteri
39	63	7.8	413	1	GAG2_HUMAN	P10264 homo sapien
40	63	7.8	417	1	KCRU_HUMAN	P12532 homo sapien
41	63	7.8	452	1	SELA_AQUAE	O67140 aquifex aeo
42	63	7.8	474	1	PNTB_HAEN	P43010 haemophilus
43	63	7.8	490	1	HDVD_CLOAM	P35792 clostridium
44	63	7.8	775	1	YTX1_XENLA	P14380 xenopus lae
45	62.5	7.7	279	1	YFES_ECOLI	P76550 escherichia

## ALIGNMENTS

RESULT 1  
ID LCTR2\_HUMAN STANDARD: PRT: 151 AA.  
AC 014960: 014565;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Leukocyte cell-derived chemotaxin 2 precursor (hLCTR2).  
GN LCTR2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_Taxid:9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANT ILE-58.  
RC TISSUE=Liver;  
RX MEDLINE=98193133; PubMed=9524238;  
RA Yamagoe S., Mizuno S., Suzuki K.;  
RT "Molecular cloning of human and bovine LCTR2 having a neutrophil chemotactic activity and its specific expression in the liver";  
RL Biochim. Biophys. Acta 1396:105-113(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peripheral blood;  
RX MEDLINE=98207247; PubMed=9545637;  
RA Yamagoe S., Kameoka Y., Hashimoto K., Mizuno S., Suzuki K.;  
RT "Molecular cloning, structural characterization, and chromosomal mapping of the human LCTR2 gene";  
RL Genomics 48:324-329(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99048897; PubMed=9832057;  
RA Nagai H., Hamada T., Uchida T., Yamagoe S., Suzuki K.;  
RT "Systemic expression of a newly recognized protein, LCTR2, in the human body";  
RL Pathol. Int. 48:882-886(1998).  
RN [4]  
RP SEQUENCE FROM N.A.  
RX Kimmey W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M., Kadner K., Miguel T., Miller C., Pittluck S., Pollard M., Rojeski H., Subramanian S., Martin C.H.;  
RT Submitted (May-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RX Strong C., Biewald T., Tin-Wollam A., Duckels G.;  
RT Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SUBCELLULAR LOCATION.  
RX MEDLINE=97031488; PubMed=8877413;  
RA Yamagoe S., Yamakawa Y., Matsuo Y., Minowada J., Mizuno S., Suzuki K.;  
RT "Purification and primary amino acid sequence of a novel neutrophil chemotactic factor LCTR2";  
RL Immunol. Lett. 52:9-13(1996).  
CC -!- FUNCTION: Has a neutrophil chemotactic activity. Also a positive regulator of chondrocyte proliferation.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted.

```

CC -I- TISSUE SPECIFICITY: Highly expressed in adult and fetal liver and
CC weakly in testis. Not expressed in bone marrow.
CC -I- INDUCTION: By phytohemagglutinin (PHA).
CC -I- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D63521; BAA23609.1; -.
DR EMBL: AB007546; BAA25669.1; -.
DR EMBL: AC004763; AAC17734.1; -.
DR EMBL: AC002428; AAB66905.1; -.
DR Genew: HGNC:6550; LECT2.
DR MIM: 602882; -.
RW Chemotaxis; Signal; Polymorphism.
FW SIGNAL 1 18 BY SIMILARITY.
FT CHAIN 19 151 LEUCOCYTE CELL-DERIVED CHEMOTAXIN 2.
FT VARIANT 58 58 V -> I.
FT SEQUENCE 151 AA; 16376 MW; A4267E0A24E69631 CRC64;
SQ /FTID=VAR_011386.
Query Match 99.9%; Score 806; DB 1; Length 151;
Best Local Similarity 99.3%; Pred. No. 2,4e-76;
Matches 150; Conservative % 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESTKALLLGLISTALAGPWANICAGKSSNEIFRTCDRHCCGGYSADRSQRPHGVDTLC 60
DB 1 MFTSKALLLGLISTALAGPWANICAGKSSNEIFRTCDRHCCGGYSADRSQRPHGVDTLC 60
QY 61 SASSTVAPFTGMIVGQEKPYQKNNAINNNGVIRISGRFCVKMPYIKRKTKKPKKGEKL 120
DB 61 SASSTVAPFTGMIVGQEKPYQKNNAINNNGVIRISGRFCVKMPYIKRKTKKPKKGEKL 120
QY 121 GTLLPLQKYPFGIQSHVHIENCSSDPAYL 151
DB 121 GTLLPLQKYPFGIQSHVHIENCSSDPAYL 151
RESULT 2
ID LECT2_BOVIN STANDARD; PRT; 151 AA.
AC 062644;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukocyte cell-derived chemotaxin 2 precursor (bLECT2) (Chondromodulin
DE II) (bcbm-II).
GN LECT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovine; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN SEQUENCE FROM N.A.
RP TISSUE=Liver;
RX MEDLINE=98193133; Pubmed=9524238;
RA Yamagoe S., Mizuno S., Suzuki K.;
RT "Molecular cloning of human and bovine LECT2 having a neutrophil
RT chemotactic activity and its specific expression in the liver.";
RL Biochim. Biophys. Acta 1396:105-113(1998).
RN [2]
RN SEQUENCE FROM N.A.
RP TISSUE=Embryo;
RX MEDLINE=99160594; Pubmed=10050029;
RA Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A.,
RA Hirai T.;
RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the

```

```

RT growth-promoting actions of bovine recombinant protein." ;
RL J Biochem. 125:436-442(1999).
RN [3]
RP SEQUENCE OF 19-15L.
RC TISSUE=Epiphyseal cartilage;
RX MEDLINE=96394331; PubMed=8798437;
RA Hiraki Y., Inoue H., Kondo J., Kamizono A., Yoshitake Y.,
Rv Shukunami C., Suzuki F.;
RT "A novel growth-promoting factor derived from fetal bovine cartilage, chondromodulin II. Purification and amino acid sequence.";
RL Biol. Chem. 271:22657-22662(1996)
CC -I- FUNCTION: Has a neutrophil chemotactic activity. Also a positive regulator of Chondrocyte proliferation.
CC CC -I- SUBCELLULAR LOCATION: Secreted.
CC CC -I- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
CC -----
CC DR EMBL: AB001350; BAA25302.1; -.
RW EMBL: D89011; BAB18616.1; -.
KW Chemicals; Signal.
FT SIGNAL 1 18
FT CHAIN 19 151 LEUKOCYNE CELL-DERIVED CHEMOTAXIN 2.
SQ SEQUENCE 151 AA; 16320 MW; 2A1DEBF5B28A3D8 CRC64;

Query Match      87.0% Score 702; DB 1; Length 151;
Best Local Similarity 86.1%; Pred. No. 1.4e-65;
Matches 130; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Oy 1 MFSTGALLAGLLSTAALGPANICACGRSSNIRTCDRHGGCGYSQRSORPHOGVDILC 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||::||
Db 1 MFSTGTLLLAALLSPALAGPWAIIICAGRSSNRIFRDGHGCCGYFTAQRNKLGHDGVLC 60

Oy 61 SACSIVAYAPTGMTVGCBEKPYQKNNAINNVAISRGFCVKMYTIPIKTAKPKKEKL 120
    | ||||||| ||||| ::|||::|||::|||::|||::|||::|||::|||::|||::|||::||
Db 61 SDGSTVYAPPTGKIIMGGEKPKNKA INNVRISGGFGCIKMFIYIPIKTKGSITKKSEKL 120

Oy 121 GTLLPLQKYVPGLDSHVHIENCDSDPAYYL 151
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 GTLLPLQKYVPGLDSHVHIENCDLSDPYYTL 151

RESULT 3
ICT2_MOUSE STANDARD; PRT; 151 AA.
ID ICT2_MOUSE AC O88803: O88804: OGQWNS: Q9Z337;
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Leukocyte cell-derived chemotaxin 2 precursor (Chondromodulin II) (Chm-II).
GN LECT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT VAL-129.
RC STRAIN=BALB/C; Tissue=Liver;
RX MEDLINE=98382586; Pubmed=9714793;
RA Yamagoe S., Watanabe T., Mizuno S., Suzuki K.;
RT "The mouse Lect2 gene: cloning of cDNA and genomic DNA, structural characterization and chromosomal localization." ;
RL Gene 216:171-178(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1);
RC STRAIN=Swiss Webster / NIH Swiss; Tissue=Embryo, and Liver;
```

[illegible]

```

RA      Ness S.A., Marknell A., Graf T. ;
RA      "The v-myb oncogene product binds to and activates the promyelocyte-
RT      specific mim-1 gene".
RL      Cell 59:1115-1125(1989).
RN      [2]
RP      PARTIAL SEQUENCE, AND SEQUENCE OF 295-299 FROM N.A.
RC      STRAIN-White Leghorn;
RX      MEDLINE=93011981; Pubmed=1397316;
RA      Yamada K., Tsuchiya M., Mishima K., Shimoyama M.;
RT      "p33, an endogenous target protein for arginine-specific ADP-
RT      ribosyltransferase in chicken polymorphonuclear leukocytes, is highly
RL      homologous to mim-1 protein (myb-induced myeloid protein-1)".
RL      FEBS Lett. 311:203-205(1992).
CC      -I- FUNCTION: NOT KNOWN.
CC      -I- SUBCELLULAR LOCATION: GRANULES OF PROMYELOCYTES.
CC      -I- INDUCTION: BY THE MYB ONCOGENE.
CC      -I- PTM: SUBSTRATE FOR ARGININE-SPECIFIC ADP-RIBOSYLTRANSFERASE.
CC      -I- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      at the European Bioinformatics Institute. There are no restrictions on
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M29448; AAA48954.1; -.
DR      EMBL; M29449; AAA48958.1; -.
DR      PIR; A33755; A33755.
DR      PIR; S29229; S29229.
KW      Repeat; Signal.
FT      SIGNAL             1      23      PROBABLE.
FT      CHAIN             24      326      MYELOID PROTEIN-1.
FT      REPEAT            28      162      1.
FT      REPEAT            177     312      2.
FT      CONFLICT           297     297      I -> Y (IN REF. 1).
SO      SEQUENCE           326 AA; 35636 MW; B8C4742EF0D3BAD3 CRC64;

Query Match                               53.7%; Score 433; DB 1; Length 326;
Best Local Similarity 50.3%; Pred. No. 2e-37;
Matches 80; Conservative 27; Mismatches 38; Indels 14; Gaps 2;

QY      7 LILAGLSTPLAG-----MANICAGKSSNETRTCDRHCCGGYSSAQRQR--P 52
DB      7 IALLSVSTAFAROMEVHPQOQGRHMAQICSGNPNRRIRGCRYGCGNGASRQGGK 66
QY      53 HGGVDILICSGASTYVAFPTGMIVGOEKPYQNKNAINNGYRISGRFCVKMFYIKPIKYK 112
DB      67 HKGVDVICTGDSIVYAFEFSGQLSGPIREFHNGNALIDGQVQSSGRCVKRWCHPIRYHG 126
QY      113 PIKKGKELGTLPLQKYRPGISQSHVHINCSSDPTAYL 151
DB      127 QIQGGQDLGMLPQKVEPGIIVSHIHVENCDSQSDPTHL 165

YRBA_HAEIN
ID      YRBA_HAEIN          STANDARD;          PRT;          475 AA.
AC      P44693;
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-JUN-1995 (Rel. 32, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Hypothetical protein H10409.
GN      H10409.
OS      Haemophilus influenzae.
OC      Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC      Haemophilus.
OX      NCB1_TaxID=727;
RX      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN-Rd / KW20 / ATCC 51907;
RX      MEDLINE=95350630; Pubmed=7542800;
```

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA Mckenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fitchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
RA Greth C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.,  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd.";  
RL Science 269:496-512(1995).  
CC -1- FUNCTION: COULD BE INVOLVED IN CELL WALL DEGRADATION OR  
CC FORMATION (BY SIMILARITY).  
CC -1- SIMILARITY: STRONG TO E. COLI YEBA  
CC -1- SIMILARITY: TO STAPHYLOCOCCUS LYSOSTAPHIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: U32724; AAC22068.1; -  
CC DR MEROPS: M37; UPW; -  
DR TIGR: H10409; -  
DR InterPro: IPR002886; Peptidase\_M37.  
DR Pfam: PF01551; Peptidase\_M37; 1.  
DR Hypothetical protein; Cell wall; Complete proteome.  
KW SEQUENCE 475 AA; 53255 MW; 143C10F92233939D CRC64;  
SQ  
Query Match 9.6%; Score 77.5; DB 1; Length 475;  
Best Local Similarity 28.7%; Pred. No. 1.7; Indels 11; Gaps 4;  
Matches 27; Conservative 15; Mismatches 41;  
QY 51 RPHGVLDILCSAGTYAFTGMIVGOEKPYONKNAINNGVRISGRFCVKNFYTKP--I 108  
DB 346 RPHKVDVSVSGCTPIYADGTIV--EKVAYVAGAGRYVALRGRGREGYQYVMHLSKSLV 403  
QY 109 KYGPKIKGKGLTLLPLKRVYPIGLOS--HVHIE 140  
DB 404 KAGOTVKKGRITAL-----SGNTGISGPHLHYE 432  
RESULT 6  
DAS\_PICAN STANDARD; PRT; 710 AA.  
ID DAS\_PICAN STANDARD; PRT; 710 AA.  
AC P06834;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Dihydroxy-acetone synthase (EC 2.2.1.3) (DHAS) (formaldehyde  
DE transketolase) (glycerone synthase).  
GN DAS.  
OS Pichia angusta (Yeast) (Hansenula polymorpha).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Pichia.  
OX NCBI\_TaxID=4905;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 34438;  
RX MEDLINE=85215670; PubMed=2987872;  
RA Jenewicz Z.A., Eckart M.R., Drewe C., Roggenkamp R.O.,  
RA Hollenberg C.P., Maat J., Ledebor A.M., Visser C., Verrips C.T.;  
RT "Cloning and characterization of the DAS gene encoding the major  
RT methanol assimilatory enzyme from the methylotrophic yeast Hansenula  
RT polymorpha.";  
RL Nucleic Acids Res. 13:3043-3062(1985).  
RN [2]  
RP REVISIONS: SEQUENCE OF 667-710 FROM N.A.  
RC STRAIN=ATCC 34438;

RX MEDLINE=93101130; PubMed=1465101;  
RA Hansen H., Didion T., Thiemann A., Veenhuis M., Roggenkamp R.O.;  
RT "targeting sequences of the two major peroxisomal proteins in the  
RT methylotrophic yeast Hansenula polymorpha.";  
RL Mol. Gen. Genet. 235:269-278(1992).  
CC -1- FUNCTION: THIS IS THE MAJOR METHANOL ASSIMILATORY ENZYME FROM  
CC THE METHYLOTROPHIC HANSENULA POLYMORPHA.  
CC CATALYTIC ACTIVITY: D-xylulose 5-phosphate + formaldehyde -  
CC glyceralddehyde 3-phosphate + glyceral.  
CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: Peroxisomal.  
CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: X02424; CAA26276.1; -  
CC DR PIR: A23009; XHOFK.  
DR HSP: P23254; ITRK.  
DR InterPro: IPR000360; Transketolase.  
DR Pfam: PF00456; transketolase; 1.  
DR Pfam: PF02779; transket-pyr; 1.  
DR PROSITE: PS00801; TRANSKETOLASE; 1.  
DR PROSITE: PS00802; TRANSKETOLASE; 2; 1.  
DR PROSITE: PS00342; MICRORODIESCTER; 1.  
KW Methanol utilization; Peroxisome; Transferrase; thiamine pyrophosphate.  
FT SITE 708 710 MICRORODY TARGETING SIGNAL (POTENTIAL).  
SQ SEQUENCE 710 AA; 78842 MW; E4424BD37CC8B0B CRC64;  
Query Match 9.0%; Score 73; DB 1; Length 710;  
Best Local Similarity 24.1%; Pred. No. 7.6; Indels 34; Gaps 7;  
Matches 33; Conservative 21; Mismatches 49;  
QY 11 GLISFALGPMWNTICAGKSNIEITCDRHGCGYSAORSORPHGVLDILCSAGTYVAPF 70  
DB 525 GLFQKAVELPFSLSLS-LRNEVL-----QYLABRAGRRNAGVILEDAENAEVQI 575  
QY 71 TGMIVGOEKPYONKNAINNGVRISGRF-----CVKMYPIPIKKGPI--KKEGL 120  
DB 576 IG--VGAEMEFADKAA-----KILGRFTRVLSIPCTRLFDEOSIGRSVLRDRGV 628  
QY 121 GTLLPLKRVYPIGLOSHV 137  
DB 629 PTVV-----VDGHV 637  
RESULT 7  
XDH\_RAT STANDARD; PRT; 1330 AA.  
ID XDH\_RAT STANDARD; PRT; 1330 AA.  
AC P22985; Q63157;  
DT 01-AUG-1991 (Rel. 19, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Xanthine dehydrogenase/oxidase [includes: Xanthine dehydrogenase  
DE (EC 1.1.1.204) (XD); Xanthine oxidase (EC 1.1.3.22) (XO) (Xanthine  
DE oxidoreductase)].  
GN XDH.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=Liver;  
RX MEDLINE=90354396; PubMed=2387845;  
RA Amaya Y., Yamazaki K.-I., Sato M., Noda K., Nishino T., Nishino T.;  
RT "Proteolytic conversion of xanthine dehydrogenase from the

RT NAD-dependent type to the O2-dependent type. Amino acid sequence of  
 RT rat liver xanthine dehydrogenase and identification of the cleavage  
 RT sites of the enzyme protein during irreversible conversion by  
 RT trypsin.";  
 RT J. Biol. Chem. 265:14170-14175(1990).  
 RN [2]  
 RP SEQUENCE OF 1-54 FROM N.A.  
 RC STRAIN-Sprague-Dawley;  
 RX MEDLINE=94268906; PubMed=8208609;  
 RA Chow C.W., Clark M., Rinaldo J., Chalkley R.;  
 RT "Identification of the rat xanthine dehydrogenase/oxidase promoter.";  
 RT Nucleic Acids Res. 22:1846-1854(1994).  
 CC -1- FUNCTION: THIS ENZYME CAN BE CONVERTED FROM THE DEHYDROGENASE FORM  
 CC (D) TO THE OXIDASE FORM (O) IRREVERSIBLY BY PROTEOLYSIS OR  
 CC REVERSIBLY THROUGH THE OXIDATION OF SULFHYDRYL GROUPS.  
 CC -1- CATALYTIC ACTIVITY: Xanthine + NAD(+) + H(2)O = urate + NADH.  
 CC -1- CATALYTIC ACTIVITY: Xanthine + H(2)O + O(2) = urate + H(2)O(2).  
 CC -1- COFACTOR: FAD, MOLIBDOPTERIN, AND TWO 2FE-2S CLUSTERS.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -1- INDUCTION: BY INTERFERON.  
 CC -1- SIMILARITY: BELONGS TO THE XANTHINE DEHYDROGENASE FAMILY.  
 CC -1- SIMILARITY: TO 2FE-2S FERRODOXINS IN THE N-TERMINAL DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: J05579; AAA42349.1; .  
 DR EMBL: U08122; AAA18869.1; .  
 DR EMBL: U08120; AAA18869.1; JOINED.  
 DR EMBL: U08121; AAA18869.1; JOINED.  
 DR PIR: A37810; A37810.  
 DR HSSP: P80457; IPIQ.  
 DR InterPro: IPR002888; 2Fe-2S bind.  
 DR InterPro: IPR000564; 2Fe2S\_ferredoxin.  
 DR InterPro: IPR000674; Aldxan\_dh\_C.  
 DR InterPro: IPR005107; CO\_dh flav\_C.  
 DR InterPro: IPR000572; Euk\_Mb\_oxred.  
 DR InterPro: IPR001041; Ferredoxin.  
 DR InterPro: IPR002346; dehydrog\_molypb.  
 DR Pfam: PF00111; fer2; 1.  
 DR Pfam: PF00941; FAD\_binding\_5; 1.  
 DR Pfam: PF01315; Ald\_xan\_dh\_C; 1.  
 DR Pfam: PF01799; fer2\_2; 1.  
 DR Pfam: PF02738; Ald\_xan\_dh\_C2; 1.  
 DR Pfam: PF03450; CO\_dh flav\_C; 1.  
 DR Prodom: PD16071; 2Fe-2S bind; 1.  
 DR PROSITE: PS00197; 2FE2S\_FERREDOXIN; 1.  
 DR PROSITE: PS00559; MOLIBDOPTERIN\_EUK; 1.  
 DR OXIDOREDUCTASE: NAD; Molybdenum; Flavo protein; FAD; Iron-sulfur.  
 KW INIT\_MET 0  
 FT METAL 36 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 42 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 47 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 FT METAL 50 IRON-SULFUR (2FE-2S) (BY SIMILARITY).  
 SQ SEQUENCE 1330 AA; 146111 MW; A3DD206B9D74E565 CRC64;  
 Query Match 9.0%; Score 73; DB 1; Length 1330;  
 Best local Similarity 28.8%; Pred. No. 15;  
 Matches 34; Conservative 12; Mismatches 42; Indels 30; Gaps 6;

RESULT 8  
 COL\_1ARATH STANDARD; PRT; 355 AA.  
 AC 050035;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Zinc finger protein CONSTANS-LIKE 1.  
 GN COL1 OR AT5G15850 OR F14P8\_230.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Landsberg erecta; TISSUE=Aerial parts;  
 RA Putterill J.J., Ledger S.E., Lee K., Robson F., Murphy G.,  
 RA Coupland G.;  
 RT "The flowering-time gene CONSTANS and homologue CONSTANS LIKE 1 exist  
 RT as a tandem repeat on chromosome 5 of Arabidopsis.";  
 RT (in) Plant Gene Register PGR97-077.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016721; PubMed=11130714;  
 RA Tsubata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,  
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,  
 RA Kohara M., Matsunoto M., Matsuno A., Muraki A., Nakayama S.,  
 RA Nakazaki N., Naturo K., Okumura S., Shino S., Takeuchi G., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
 RA Huang E., Spiegel L., Gao J., O'Shaughnessy A., Preston R.,  
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,  
 RA Stoneking T., Pepin K., Spielth J., Sekhon M., Armstrong J., Becker M.,  
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,  
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,  
 RA Leonard S., Meyer R., Mlvaney E., Ozersky P., Riley A., Stromwallt C.,  
 RA Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedha N.,  
 RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,  
 RA Kirchhoff K., Toth K., King L., Bahner A., Miller B., Matra M.,  
 RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,  
 RA Volckaert G., Wambutt R., Duesterhoeft A., Stekema W., Pohl T.,  
 RA Ertlan K.-D., Terry N., Hartley N., Bent E., Johnson S.,  
 RA Langham S.-A., McCullagh B., Robben J., Gymnopoulos B., Zimmermann W.,  
 RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,  
 RA van Staveren M., Dirkse W., Mooljman P., Klein Lankhorst R.,  
 RA Weltzienegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,  
 RA Feldpausch M., Lamberth S., Villarroel R., Giesen J., Ardiles W.,  
 RA Bents O., Lemcke K., Kolesov G., Mayer K., Rued S., Schoof H.,  
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.;  
 RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 408:823-826(2000).  
 RN [3]  
 RP CHARACTERIZATION.  
 RC STRAIN=cv. C24;  
 RX MEDLINE=21260187; PubMed=11359606;  
 RA Ledger S., Strayer C., Ashton F., Kay S.A., Putterill J.,  
 RT "Analysis of the function of two circadian-regulated CONSTANS-LIKE  
 RT genes.";  
 RT Plant J. 26:15-22(2001).  
 CC -1- FUNCTION: Putative transcription factor that may be involved in  
 CC the light input to the circadian clock but does not affect  
 CC flowering time.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LEAVES AND AT LOWER LEVELS  
 CC IN STEMS, FLOWERS AND STILICES, NOT DETECTED IN ROOTS.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT.  
 CC -1- INDUCTION: EXPRESSED WITH A CIRCADIAN RHYTHM SHOWING A PEAK  
 CC AT DAWN.  
 CC -1- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.

CC -1- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Y10555; CAA71587.1; -;  
 DR EMBL: Y10556; CAA71588.1; -;  
 DR EMBL: AL391144; CAC01784.1; -;  
 DR InterPro: IPR000315; znf\_Box.  
 DR InterPro: IPR002926; znf\_constants.  
 DR Pfam: PF00643; zf-B\_box; 2.  
 DR ProDom: PD007661; znf\_constants; 2.  
 DR SMART: SM00336; BBOX; 1.  
 DR PROSITE: PS01119; ZF-BBOX; 2.  
 KW Transcription regulation; Zinc-finger; DNA-binding; Nuclear protein;  
 KW Repeat; Multigene family.  
 FT ZN\_FING 12 54 B BOX-TYPE 1.  
 FT ZN\_FING 55 97 B BOX-TYPE 2.  
 FT ZN\_FING 355 AA; 39491 MW; 412C237A9CD30309 CRC64;  
 SQ SEQUENCE  
 Query Match 9.0%; Score 72.5; DB 1; Length 355;  
 Best Local Similarity 24.4%; Pred. No. 4.1;  
 Matches 30; Conservative 15; Mismatches 43; Indels 35; Gaps 4;  
 QY 20 PANNICAGKSSNEIRTCDRHCGGQYSAGRSQRPHGVDIL-----CSAGSTV 66  
 DB 62 PAAPFCKADASLCTTCDSE---PYONKNAINGVYISGRCFCVKMEYIKPIKYGK 118  
 QY 67 YAPFTGMIVGOEK-----PYONKNAINGVYISGRCFCVKMEYIKPIKYGK 113  
 DB 119 TDPENRLVVGQEEDEDEAFAASMLPNSGKSGNN-----NGSFIGDEFLNVDSSS 172  
 QY 114 IKR 116  
 DB 173 DKQ 175  
 RESULT 9  
 TRILL\_MOUSE STANDARD; PRT; 625 AA.  
 ID TRILL\_MOUSE  
 AC O35305; Q8VCT7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor receptor superfamily member 11A precursor  
 DE (Receptor activator of NF-kB) (Osteoclast differentiation factor  
 DE receptor) (ODPR).  
 GN TNFRSF11A OR RANK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Retal Liver;  
 RX MEDLINE=98032977; PubMed=9367155;  
 RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougan W.C.,  
 RA Tometsko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,  
 RA Galibert L.;  
 RA "A homologue of the TNF receptor and its ligand enhance T-cell growth  
 RA and dendritic-cell function.";  
 RL Nature 390:175-179(1997).  
 RU Nature 390:175-179(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RA Strausberg R.;  
 RU Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]

RP FUNCTION.  
 RX MEDLINE=9097247; PubMed=9878548;  
 RA Nakagawa N., Kinoshita M., Yamaguchi K., Shima N., Yasuda H., Yano K.,  
 RA Morinaga T., Higashio K.;  
 RT "RANK is the essential signaling receptor for osteoclast  
 RT differentiation factor in osteoclastogenesis.";  
 RL Biochem. Biophys. Res. Commun. 233:395-400(1998).  
 CC -1- FUNCTION: Receptor for TNFRSF11/RANKL/TRANCE/OPGL; essential for  
 CC RANKL-mediated osteoclastogenesis. Involved in the regulation of  
 CC interactions between T-cells and dendritic cells.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -1- TISSUE SPECIFICITY: UNIDITOUS EXPRESSION WITH HIGH LEVELS IN  
 CC TRABECULAR BONE, THYMUS, SMALL INTESTINE, LUNG, BRAIN AND KIDNEY.  
 CC WEAKLY EXPRESSED IN SPLEEN AND BONE MARROW.  
 CC -1- SIMILARITY: CONTAINS 4 TNFR-CYS REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF019046; AAB86810.1; -;  
 DR EMBL: BC019185; AAH19185.1; -;  
 DR HSSP: P25942; ICDF.  
 DR MGD: MGI:1314891; Tnfrsf11a.  
 DR InterPro: IPR001368; TNFR\_c6.  
 DR Pfam: PF00020; TNFR\_c6; 3.  
 DR ProDom: PD000771; TNFR\_c6; 1.  
 DR SMART: SM00208; TNFR; 4.  
 DR SMART: SM00208; TNFR; 4.  
 DR PROSITE: PS00652; TNFR\_NGFR\_1; 1.  
 DR PROSITE: PS00650; TNFR\_NGFR\_2; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Repeat; signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 625  
 FT DOMAIN 31 214  
 FT TRANSMEM 215 234  
 FT DOMAIN 235 625  
 FT REPEAT 35 69  
 FT REPEAT 72 113  
 FT REPEAT 115 152  
 FT REPEAT 155 195  
 FT DISULFID 35 47  
 FT DISULFID 48 61  
 FT DISULFID 51 69  
 FT DISULFID 72 87  
 FT DISULFID 93 113  
 FT DISULFID 115 128  
 FT DISULFID 134 152  
 FT CARBOHYD 106 106  
 FT CARBOHYD 175 175  
 FT CONFLICT 494 494  
 SQ SEQUENCE 625 AA; 66621 MW; F8C1872B95511D8E CRC64;  
 Query Match 8.7%; Score 70.5; DB 1; Length 625;  
 Best Local Similarity 21.4%; Pred. No. 12;  
 Matches 31; Conservative 12; Mismatches 39; Indels 63; Gaps 5;  
 QY 16 ALAGPWANL-----CAGKSSNEIRTCDRHCGGQYSAGRSQRPHGVDILCSA 62  
 DB 309 AAGGPAEVRDSRTFLVSEVETGDLRSKIPTEDEY-----TDRPSQPSSTLSLIQ 362  
 QY 63 GSIVYAPFTGMIVGOEKPYONKNAINGVYISGRCFCVKMEYIKPIKYGKGLGT 122  
 DB 363 GSKSIPF-----QE-----PLEVGENDS- 381  
 QY 123 LPLQKVPYPGIOSVHIENCSDSDP 147  
 DB 382 ---LSQCTGTSTVDSGCDFTPEP 403





RT "Imprinting and evolution of two Kruppel-type zinc-finger genes, ZIM3  
 RT and ZNE264, located in the PEG3/USP29 imprinted domain.";  
 CC Genomics 77:91-98(2001).  
 CC -1- FUNCTION: May function as a transcription factor.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF365931; AAL11635.1; -;  
 DR Genbank: HGNC:16366; ZIM3.  
 DR InterPro: IPR001909; KRAB.  
 DR InterPro: IPR000822; Znf\_C2H2.  
 DR Pfam: PF00096; zf\_C2H2; 1.  
 DR Pfam: PF01352; KRAB; 1.  
 DR PROSITE: PS50805; KRAB; 1.  
 DR PROSITE: PS50028; ZINC\_FINGER\_C2H2\_1; 1.  
 DR PROSITE: PS50157; ZINC\_FINGER\_C2H2\_2; 1.  
 DR Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;  
 KW DNA-binding; Repeat.  
 FT DOMAIN 8 80 KRAB.  
 FT ZINC\_FINGERS.  
 FT ZN\_FING 167 470  
 FT ZN\_FING 167 189  
 FT ZN\_FING 195 217 C2H2-TYPE.  
 FT ZN\_FING 223 245 C2H2-TYPE.  
 FT ZN\_FING 251 273 C2H2-TYPE.  
 FT ZN\_FING 279 301 C2H2-TYPE.  
 FT ZN\_FING 307 329 C2H2-TYPE.  
 FT ZN\_FING 335 357 C2H2-TYPE.  
 FT ZN\_FING 363 385 C2H2-TYPE.  
 FT ZN\_FING 391 413 C2H2-TYPE.  
 FT ZN\_FING 419 441 C2H2-TYPE.  
 FT ZN\_FING 447 470 C2H2-TYPE.  
 SO SEQUENCE 472 AA; 54498 MW; 755279B78653F286 CRC64;  
 Query Match 8.6%; Score 69.5; DB 1; Length 472;  
 Best Local Similarity 24.1%; Pred. No. 11;  
 Matches 34; Conservative 19; Mismatches 67; Indels 21; Gaps 5;  
 QY 1 MSTRALLAGLSTALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRORPHOGVDILC 60  
 DB 175 LFSKSRSLQSHLRHACQKPFECSCGRAFEKMKLKH---QKTHAEPRPK-----C 225  
 QY 61 SAGSYVAEFTGMIVG-----EKPYONKNAIINGVRISGRFCYKMFYIKPIKGPPIK 115  
 DB 226 EMCGNAYKOKSNLFOHOKMTKEKPYOCKTC---GKAFSWKSSCINHEKIHNAKKSYQCN 282  
 QY 116 KGEKL---GTLPLQKVYPG 132  
 DB 283 ECEKSPFONSTLIQHKVHTG 303  
 RESULT 12  
 TRF3\_THEVO STANDARD; PRT; 779 AA.  
 ID TRF3\_THEVO  
 AC Q97AJ6;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tricorn protease interacting factor F3 (EC 3.4.11.-).  
 GN TRF3 OR TV0814 OR TVG0817891.  
 OS Thermoplasma volcanium.  
 CC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 CC Thermoplasmataceae; Thermoplasma.  
 CC NCBI\_TaxID=50339;  
 RP

EN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GS51 / DSM 4299 / JCM 9571;  
 RX MEDLINE=20570466; PubMed=11121031;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima T., Yamamoto Y., Aramaki H., Makino K., Kawamoto T.,  
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 RT sequence of Thermoplasma volcanium";  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 CC -1- FUNCTION: Proteases F1, F2 and F3 degrade oligopeptides produced  
 CC by tricorn (themselves probably produced by the proteasome),  
 CC yielding free amino acids (by similarity).  
 CC -1- COFACTOR: Binds one zinc ion (by similarity).  
 CC -1- SUBUNIT: Part of the tricorn proteolytic complex (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP000993; BAB59956.1; -;  
 DR MEROPS: M01.021; -;  
 DR InterPro: IPR001930; Ala-peptase.  
 DR InterPro: IPR000130; Zn\_MTPeptidse.  
 DR Pfam: PF01433; Peptidase\_M1; 1.  
 DR PRINTS: PR00756; ALADIPTASE.  
 DR PROSITE: PS00142; ZINC\_PROTEASE; 1.  
 KW Hydrolyse; Metalloprotease; Aminopeptidase; Zinc; Complete proteome.  
 FT METAL 266 266 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 267 267 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 270 270 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 289 289 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 352 352 PROTON DONOR (POTENTIAL).  
 SO SEQUENCE 779 AA; 88068 MW; 6BC227F2ECCB201 CRC64;  
 Query Match 8.6%; Score 69; DB 1; Length 779;  
 Best Local Similarity 26.4%; Pred. No. 22;  
 Matches 33; Conservative 21; Mismatches 59; Indels 12; Gaps 4;  
 QY 24 ICAGKSSNEIRTCDRHGGCGQYSAQRORPHOGVDILCSAGSYVAEFTGMIVGQEKPYON 83  
 DB 590 IAVGKASRLALLDEAVC-ETLAPRFSNFEQTELEKSAIVAVLSTGVDGVMVKYS 648  
 QY 84 KNAIINGVR-ISGRG---FCVKMFYIKPIKPKKKEKLG-----TLPLQKVYPG 132  
 DB 649 LRDDEKVKLIISGFKLSTSLDLSVSGMIEKGIKODMLSFYLSALETMGAREYISN 708  
 QY 133 IQSHV 137  
 DB 709 LENIV 713  
 RESULT 13  
 YT63\_CAEL STANDARD; PRT; 184 AA.  
 ID YT63\_CAEL  
 AC Q11079;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 20.9 kDa protein B0563.3 in chromosome X.  
 GN B0563.3.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 CC NCBI\_TaxID=6239;  
 CC [1]  
 RP SEQUENCE FROM N.A.

CC STRAIN-Bristol N2;  
 RA Favello T.;  
 RL Submitted (JUN-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE BIL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U28740; AAA68321.1; -  
 DR Wormpep: B0563.3; CE02441.  
 DR Interpro: IPR002199; Bax\_inhbrl1.  
 DR Pfam: PF01027; UPF0005; 1.  
 KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 70 90 POTENTIAL.  
 FT TRANSMEM 96 116 POTENTIAL.  
 FT TRANSMEM 123 143 POTENTIAL.  
 FT TRANSMEM 160 180 POTENTIAL.  
 SQ SEQUENCE 184 AA; 20942 MW; 03836D218202F378 CRC64;  
 Query Match 8.4%; Score 68; DB 1; Length 184;  
 Best Local Similarity 27.2%; Pred. No. 5.9;  
 Matches 31; Conservative 9; Mismatches 32; Indels 42; Gaps 7;  
 QY 25 CAKSSNEIRTCDRHGGCGQYSAQ--RSQRPHGVLDLCSAGS-----TVY--- 67  
 Db 21 CVPKRS-----RFGNSDYCAPRLRQSPSQ---LCSFGRITRSCCHYGMULTEPAK 69  
 QY 68 -AFTGMIYGOEKPY--QNKNAINNGVRIISGRFCV-----KMFYIKP 107  
 Db 70 VLEAAVITGLVNASLRAVYTIQNKRPDSVGASMGSLCYLLMAGITQMFMPSP 123  
 RESULT 14  
 MUS2\_AOUAE STANDARD; PRT; 762 AA.  
 ID MUS2\_AOUAE  
 AC 067287;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE MutS2 protein.  
 GN MUTS2 OR MUTS2 OR AQ\_1242.  
 OS Aquifex aeolicus.  
 CC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;  
 CC Aquifex.  
 CC NCBI\_TaxID=63363;  
 RX MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Snead M.A., Keller M., Anjay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;  
 RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus";  
 RL Nature 392:353-358(1998).  
 CC -1- FUNCTION: NOT KNOWN.  
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTS FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE00730; AAC07247.1; -

DR Interpro: IPR000302; KOW\_motif.  
 DR Interpro: IPR000432; MutS\_C.  
 DR Interpro: IPR002863; MutS\_N.  
 DR Interpro: IPR002625; Smr/MutS2.  
 DR Pfam: PF00467; KOW; 1.  
 DR Pfam: PF00488; MutS\_C; 1.  
 DR Pfam: PF01713; Smr; 1.  
 DR Prodom: PD001263; MutS\_C; 1.  
 DR SMART: SM00534; MutS2; 1.  
 DR SMART: SM00533; MutS2; 1.  
 DR SMART: SM00463; Smr; 1.  
 DR TIGRfams: TIGR01069; mutS2; 1.  
 DR PROSITE: PS00486; DNA\_MISMATCH\_REPAIR\_2; FALSE\_NEG.  
 KW ATP-binding; DNA-binding; Complete proteome.  
 FT NP\_BIND 329 336 ATP (POTENTIAL).  
 SQ SEQUENCE 762 AA; 87538 MW; 3A7D977DC2EB01BC CRC64;  
 Query Match 8.4%; Score 68; DB 1; Length 762;  
 Best Local Similarity 21.8%; Pred. No. 27;  
 Matches 27; Conservative 19; Mismatches 40; Indels 38; Gaps 4;  
 QY 64 STYVAPFTGMVGOEKPYQNKNAINNGVRIISGRF-----CYKMYIRPIKPKPKKGEK 119  
 Db 20 SYASPAFTKEKTIQNLKPYTNKEVKEIELSAFTIAENVRLEFEEDIREL--LTKAKL 77  
 QY 120 LGTL-----PLQKYPGQISHVHLEN--CDSSDP 147  
 Db 78 QGALGVEDILKILNVILTKRIRRVLSHVQRLEPLRKYYKKLTSPLENLIGSIDP 137  
 QY 148 TAYL 151  
 Db 138 RGFV 141  
 RESULT 15  
 FA7\_RABIT STANDARD; PRT; 444 AA.  
 ID FA7\_RABIT  
 AC P98139; P79224;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Coagulation factor VII precursor (EC 3.4.21.21) (Serum prothrombin  
 DE conversion accelerator).  
 GN F7.  
 OS Oryctolagus cuniculus (Rabbit).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 CC NCBI\_TaxID=9986;  
 RX MEDLINE=93190306; PubMed=833365;  
 RA Brothers A.B., Clarke B.J., Sheffield W.P., Blajchman M.A.;  
 RT "Complete nucleotide sequence of the cDNA encoding rabbit coagulation  
 RT factor VII";  
 RL Thromb. Res. Suppl. 69:231-238(1993).  
 CC [2]  
 CC REVISION TO 395.  
 CC TISSUE=Liver;  
 RA Ruiz S.R., Blajchman M.A., Clarke B.J.;  
 RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: CIRCULATES IN THE BLOOD IN A ZYMOGEN FORM. FACTOR VII IS  
 CC CONVERTED TO FACTOR VIIA BY FACTOR XIa, FACTOR XIIa, FACTOR IXa, OR  
 CC THROMBIN BY MINOR PROTEOLYSIS. IN THE PRESENCE OF TISSUE FACTOR  
 CC AND CALCIUM IONS, FACTOR VIIa THEN CONVERTS FACTOR X TO FACTOR Xa  
 CC BY LIMITED PROTEOLYSIS. FACTOR VIIa WILL ALSO CONVERT FACTOR IX TO  
 CC FACTOR IXa IN THE PRESENCE OF TISSUE FACTOR AND CALCIUM (BY  
 CC SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-|-Ile bond in factor X to  
 CC form factor Xa.  
 CC -1- SUBUNIT: HETERODIMER OF A LIGHT CHAIN AND A HEAVY CHAIN LINKED  
 CC BY A DISULFIDE BOND (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: PLASMA.

CC -1- PTH: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME  
 CC GLUTAMIC ACID RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND  
 CC CALCIUM (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.  
 CC -1- SIMILARITY: CONTRAINS 2 EGF-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U77477; AAB37326.1; -  
 DR HSSP: P08709; IFAK.  
 DR MEROPS: S01.215; -  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001314; Chymotrypsin.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR002383; GLA\_blood.  
 DR InterPro: IPR001254; Ser\_protease\_Try.  
 DR InterPro: IPR000294; Vitk\_dep\_GLA.  
 DR Pfam: PF00008; EGF; 2.  
 DR Pfam: PF00089; trypsin; 1.  
 DR Pfam: PF00594; gla; 1.  
 DR PRINTS: PR00722; CHYMOTRYPSIN.  
 DR PRINTS: PR00001; GLABLOOD.  
 DR SMART: SM00179; EGF\_CA; 1.  
 DR SMART: SM00001; EGF\_like; 1.  
 DR SMART: SM00069; GLA; 1.  
 DR SMART: SM00020; Tryp\_Spc; 1.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS01187; EGF\_CA; 1.  
 DR PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
 DR Hydroxylase; Serine protease; Blood coagulation; Zymogen; Glycoprotein;  
 KW Liver; Plasma; Vitamin K; Calcium-binding; Gamma-carboxyglutamic acid;  
 KW EGF-like domain; Repeat; Signal; Hydroxylation.  
 FT SIGNAL 1 21  
 FT PROPEP 22 39  
 FT CHAIN 40 191  
 FT CHAIN 192 444  
 FT DOMAIN 45 74  
 FT DOMAIN 85 121  
 FT DOMAIN 126 167  
 FT DOMAIN 192 444  
 FT SITE 191 192  
 FT  
 FT ACT\_SITE 232 232  
 FT ACT\_SITE 281 281  
 FT ACT\_SITE 383 383  
 FT BINDING 377 377  
 FT BINDING 56 61  
 FT DISULFID 89 100  
 FT DISULFID 94 109  
 FT DISULFID 111 120  
 FT DISULFID 130 141  
 FT DISULFID 137 151  
 FT DISULFID 153 166  
 FT DISULFID 174 301  
 FT DISULFID 198 203  
 FT DISULFID 217 233  
 FT DISULFID 349 368  
 FT DISULFID 379 407  
 FT MOD\_RES 45 45  
 FT MOD\_RES 46 46  
 FT MOD\_RES 46 46  
 GAMMA-CARBOXYGLUTAMIC ACID.  
 GAMMA-CARBOXYGLUTAMIC ACID.  
 GAMMA-CARBOXYGLUTAMIC ACID.

FT MOD\_RES 53 53  
 FT MOD\_RES 55 55  
 FT MOD\_RES 58 58  
 FT MOD\_RES 59 59  
 FT MOD\_RES 64 64  
 FT MOD\_RES 65 65  
 FT MOD\_RES 68 68  
 FT MOD\_RES 74 74  
 FT MOD\_RES 102 102  
 FT MOD\_RES 111 111  
 FT CARBOHYD 211 211  
 FT CARBOHYD 242 242  
 FT CARBOHYD 306 306  
 SQ SEQUENCE 444 AA; 49011 MW; 0481ABC4FE5427F8 CRC64;

Query Match 8.4%; Score 67.5; DB 1; Length 444;  
 Best Local Similarity 27.4%; Pred. No. 17;  
 Matches 26; Conservative 11; Mismatches 39; Indels 19; Gaps 3;  
 QY 24 ICAGKSSNEIRICDRH-----GGGQY-----SAQRSQRPFGVDILCSAGSTV 66  
 Db 108 ICFCLEDEGRNCEKKNKDQLICMYEGGCEQYCSHDVGSQRSCHGEGYTLPLPGVSC- 166  
 QY 67 YAPFTGMIVGQEKPYONKNAINNGVRIISGRCVCK 101  
 Db 167 -TPVDYPCGKVPALKKRGASNPQGRIVGKVCCK 200

Search completed: March 25, 2003, 17:20:06  
 Job time : 9.5 secs



```

Q9DFJ1 ID Q9DFJ1 PRELIMINARY; PRT; 156 AA.
AC Q9DFJ1;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)
DE Chemotaxin (Fragment).
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCB1_TaxID=8022;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;
RT "Immune-relevant (including acute phase) genes identified in the
RT livers of rainbow trout, Oncorhynchus mykiss, by means of suppression
RT subtractive hybridization.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF271114; AAG28030.1;
FT NON_TER 156
SO SEQUENCE 156 AA; 17087 MW; B19067665FDA70DA CRC64;

Query Match 46.5%; Score 375; DB 13; Length 156;
Best Local Similarity 47.3%; Pred. No. 3.4e-32;
Matches 69; Conservative 29; Mismatches 46; Indels 2; Gaps 1;

QY 7 LLLAGLSTRALAGPMANICGKSSNEIRTCDRHCGCGQSQRPHGVDILCSGSIY 66
DB 10 VLLIVLSECEMYKFSQCLCDNSNRRRTGRWQGHGASRGRAHGLDITVCNDGATV 69
QY 67 YAPFTGMIVGOEKPYON--NNAINNGVRISGRFCVKMFYIKPIKYGKIGKIGTLL 124
DB 70 YAPFTVKNLNGKIVYVTPDKKAATINDGINTLSEGLCFKLFYKPKDYSGVYKNGGRIGTML 129
QY 125 PLQKVPYGIOSHVHENCSSDPTAY 150
DB 130 TWQSYPGITSHVYQMCDSKSDPTKF 155

RESULT 3
Q21241 ID Q21241 PRELIMINARY; PRT; 472 AA.
AC Q21241;
DT 01-NOV-1996 (TREMblrel. 01, Created)
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE Hypothetical 53.3 kDa protein.
GN K05F1.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdillidae; Rhabditidae;
OC Rhabdillidae; Peloderinae; Caenorhabditis.
OX NCB1_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Wohldmann P.;
RT "The sequence of C. elegans cosmid K05F1.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";

```

```

RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U29377; AAA68720.2;
KW Hypothetical protein.
SQ SEQUENCE 472 AA; 53330 MW; D7A8D9666A7190C2A CRC64;

Query Match 18.2%; Score 146.5; DB 5; Length 472;
Best Local Similarity 27.1%; Pred. No. 3.1e-07;
Matches 45; Conservative 28; Mismatches 62; Indels 31; Gaps 9;

QY 7 LLLAGL-ISTALGPMANICAGKSSNEIRTC-DRHG-CGQYSQR-SQRPHGVDILCSA 62
DB 146 LTLIGTIGGQTLGCKMKKVCNNAKNDFRCPTDSCGNTERTSGEITDGVDRCHL 205
QY 63 GSTVYAPFTGMIVGOEKPY---QNKNAINNGVRISG---RQFCVKMFYIKPIKYGPIK 115
DB 206 GEPITAPRIGEMVFW-RPYGGRKRSCADGVRLEGTOQMVGVAHISVKLSFEGHVE 264
QY 116 KKEKIGTLLPL-----QKVPYGIQSHV-----HIENC 142
DB 265 AGDEIGELNRYCFNDRGQNDVEPHEIRLYKEGRLLDPTHTLQNC 310

RESULT 4
001719 ID 001719 PRELIMINARY; PRT; 689 AA.
AC 001719;
DT 01-JUL-1997 (TREMblrel. 04, Created)
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Myb-related transcription factor (Fragment).
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinodermata; Echinodermata; Echinodermata; Echinodermata;
OC Strongylocentrotus.
OX NCB1_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98088684; PubMed=9428408;
RA Coffman J.A., Kirchhamer C.V., Harrington M.G., Davidson E.H.;
RT "SPMYb functions as an intramodular repressor to regulate spatial
RT expression of Cyttin in sea urchin embryos.";
RL Development 124:4717-4727(1997).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 3 MYB-LIKE DOMAINS.
DR EMBL; U96090; AAC47807.1;
DR HSSP; P06876; IMR.
DR InterPro; IPR001005; Myb_DNA_binding.
DR Pfam; PF00249; MYB_DNA_binding; 3.
DR SMART; SM00395; SANT; 3.
DR PROSITE; PS00037; MYB_1; UNKNOWN_3.
DR PROSITE; PS00334; MYB_2; 3.
DR PROSITE; PS50090; MYB_3; 3.
KW DNA-binding; Nuclear protein.
FT NON_TER 1
SQ SEQUENCE 689 AA; 77241 MW; 0FASABD7AF8766C6 CRC64;

Query Match 10.7%; Score 86.5; DB 5; Length 689;
Best Local Similarity 21.4%; Pred. No. 1.2;
Matches 36; Conservative 22; Mismatches 39; Indels 71; Gaps 9;

QY 17 LAGPMANIC--AGKSSNEIR----TCDR--HCGQYSQR-----SORPHQ 54
DB 163 LGRNWAETAKLLPGRTDAIKLHMNSYMKRKVETCNPTPTPKTTKTPHYHTNDQKP-- 220
QY 55 GVDILCSAGSIYAPFTGMIVGOEKPYQNKNAINNGVRIISGRFCVKMFYIKPIKYGPI 114
DB 221 -----CSSSSKITYP-----DSDFMANNSIRDALMRQGGQRYV----- 253
QY 115 KKEKIGTLLPL-----QKVPYGI-----QSHVHENCSSDP 147
DB 254 -----VRLTYPMGHDTLGGDEGSGKYKPGVKTPOKMLIMNCGEISIP 296

```

RESULT 5  
Q96ZEB PRELIMINARY; PRT; 272 AA.  
AC Q96ZEB; 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Hypothetical protein S1885.  
GN S1885.  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
NCBI\_TaxID=111955;  
RN NCBI\_TaxID=111955;  
RP SEQUENCE FROM N.A.  
RC STRAIN-JCM 10545 / 7;  
RX PubMed=11572479;  
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Ankal A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kuchida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yanagis M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT \*Complete genome sequence of an aerobic thermophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain7.\*;  
RL DNA Res. 8:123-140(2001).  
DR EMBL; AF000988; BAB66977.1; .  
DR InterPro; IPR000515; BPD.transp.  
DR PROSITE; PS00402; BPD.TRANSF.INN.MEMBR; UNKNOWN\_1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 272 AA; 30871 MW; 810C97413773BF05 CRC64;

Query Match 10.5%; Score 85; DB 17; Length 272;  
Best Local Similarity 26.9%; Pred. No. 0.6;  
Matches 28; Conservative 21; Mismatches 43; Indels 12; Gaps 5;

OY 54 QGVDIICASGVYAPFTGMT-----VGDEKPYQNKAINNGVRISGRCVQKMYIK 106  
DB 23 KAIDSSPDQENFSPSGIIEKIEKIRGNKY-TKIDYDVVMYINSGRIKVLHVE 81  
OY 107 PIYKKG-PIKKGKLTLLPLQKY-PGIQSHVHENCSSDPT 148  
DB 82 PYLEGSEIKEGEKIGKF--LESPYAGDFKHAHIEGITFEKPS 123  
RESULT 6  
Q9PTR2 PRELIMINARY; PRT; 498 AA.  
AC Q9PTR2; 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Cytochrome P450 2P2.  
GN CYP2P2.  
OS Fundulus heteroclitus (Killifish) (Mummichog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;  
OC Cyprinodontiformes; Fundulidae; Fundulus.  
NCBI\_TaxID=8078;  
RN NCBI\_TaxID=8078;  
RP SEQUENCE FROM N.A.  
RC TISSUE-LIVER;  
RA Oleksiak M.F., Stegeman J.J.;  
RT \*Diversity of vertebrate cytochrome P450 2 family genes: Cloning of  
RT three new genes and phylogenetic analysis of CYP2 gene subfamilies in  
RT Fundulus heteroclitus.\*;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-LIVER;  
RA Oleksiak M.F., Wu S., Parker C., Zeldin D.C., Stegeman J.J.;  
RT "Cloning and expression of a novel teleost cytochrome P450, CYP2P3;

RT conservation of arachidonic acid epoxigenase/19-Hydroxylase.\*;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL; AF117342; AAF21999.1; .  
DR HSSP; P00179; 1PT6.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 498 AA; 57340 MW; 1FF4ED84BBD610C9 CRC64;

Query Match 10.0%; Score 81; DB 13; Length 498;  
Best Local Similarity 21.7%; Pred. No. 3.2;  
Matches 40; Conservative 29; Mismatches 57; Indels 58; Gaps 10;

OY 1 MFSTKALLAGLISATLGPW-----ANTAGKSSNERTCDRHCGGYSAQS 49  
DB 117 VFKKKGIVMSN-----GYPMVQRRFALHALNFGLGKRTME-----RYIQEC 160  
OY 50 Q-----RPHQGVDIICASGVYAPFTGMTVGDEKPYQNK-----NAINNGVR 92  
DB 161 QYLNVEVVDQGGKPPSGGTLINNAVSNIT---ICCLVFGNRFEDDKRYHTLIDMNNELLR 217  
OY 93 ISGRGFCVQKMYIKPIKYY--GPIKKGKLTLLPLQKYPGIQSHV--HIENCSSDP 147  
DB 218 LQG-GFWVQVYVNFPSVKKMLPGPHKK-----IFIHQKTIIDFLRIKIRENDPESP 271  
OY 148 TAYL 151  
DB 272 RDTY 275

RESULT 7  
Q9PGX9 PRELIMINARY; PRT; 417 AA.  
AC Q9PGX9; 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE Hypothetical protein Xf0167.  
GN Xf0167.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
NCBI\_TaxID=2371;  
RN NCBI\_TaxID=2371;  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Agencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Farto J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Medeira A.M.B.N., Medeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Wenck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto R.B., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,

RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Senubal J.C.;  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*."  
 DR Nature 406:151-159(2000).  
 RL EMBL; AE003870; AAF82980.1; -.  
 DR InterPro; IPR002886; Peptidase\_M37.  
 DR Pfam; PF01551; Peptidase\_M37.1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 417 AA; 46554 MW; 5483743CD61061PD CRC64;

Query Match 10.0%; Score 80.5; DB 16; Length 417;  
 Best Local Similarity 25.0%; Pred. No. 3;  
 Matches 29; Conservative 24; Mismatches 42; Indels 21; Gaps 7;

QY 34 RUCDHGCGQYSAQRORPHQGVILCSAGSTVYAPFTGMI--VGQEKPYONKNAINNGV 91  
 DB 249 RLSSTFGRLRRHPLGTMRMHKGVDAATGTPIMAGDARVOFTQQRGYNVILNH-- 306  
 QY 92 RISGRGFCVYKMFYIKPIKYGPIKKGKEL--GTLPLQKVPYGIQS----HVHIE 140  
 DB 307 ---GKGY--RTLYGHMSRF--GKIKAGCKINGTYIG---YVGMGTGLATGPHLHYE 352

## RESULT 8

Q98088 PRELIMINARY; PRT; 750 AA.  
 AC 098088;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE LIPOPROTEIN.  
 GN MYPV\_4780.  
 OS Mycoplasma pulmonis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2107;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=UAB CTIP;  
 RX MEDLINE=21267165; PubMed=11353084;  
 RA Chambaud I., Hellig R., Ferris S., Barbe V., Samson D., Gallison F.,  
 RA Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,  
 RA Blanchard A.;  
 RT "The complete genome sequence of the murine respiratory pathogen  
 RT Mycoplasma pulmonis."  
 RL Nucleic Acids Res. 29:2145-2153(2001).  
 DR EMBL; AL45564; CAC13651.1; -.  
 DR Mypulist; MYPV\_4780; -.  
 KW Complete proteome.  
 SQ SEQUENCE 750 AA; 86325 MW; 37F06FE161602D5 CRC64;

Query Match 9.8%; Score 79; DB 16; Length 750;  
 Best Local Similarity 22.8%; Pred. No. 8.5;  
 Matches 42; Conservative 25; Mismatches 55; Indels 62; Gaps 8;

QY 12 LISTLAGPMANTCAGKSSNEIRTCDR--HCGQYSAQRORPHQGVILCSAGSTVYAP 69  
 DB 456 IQSPIYSGWMLD-----DRSIVTSKASDKNTKDILHGEDILLIQNKREVIAP 502  
 QY 70 FTGMIYGO--EKPYQ-----NKNAINNGVRIISGRGFCVYMF-- 103  
 DB 503 FDGKITASYAPSPYAYAGLVITLVEKMKDLVGQIDQSVIDNOILAETDRITYIAFMHJN 562  
 QY 104 -----YIKPIKKG-----PIKKGKELGTLPLQKVPYGIQSHVHIE--NCD 143  
 DB 563 PSYLENYGLVEQSSSTAIEITPATPKYAKGEVIG--LVGEFKNGGMPVHIEVSLG 621  
 QY 144 SSDP 147  
 DB 622 STNP 625

## RESULT 9

Q9UF93 PRELIMINARY; PRT; 651 AA.  
 AC 09UF93;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE Hypothetical 70.4 kDa protein (fragment).  
 GN DKF2P566N1047.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=KIDNEY.  
 RA Ottewill B., Obermaier B., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL13109; CAB61414.1; -.  
 DR InterPro; IPR002965; P\_rich\_extensn.  
 DR Pfam; PF01424; R3H; 1.  
 DR PRINTS; PR01217; PRICHEXTENS.  
 DR SMART; SM00393; R3H; 1.  
 KW Hypothetical protein.  
 FT NON\_TER 1  
 SQ SEQUENCE 651 AA; 70424 MW; 4A8DF18AB50AD04C CRC64;

Query Match 9.4%; Score 76; DB 4; Length 651;  
 Best Local Similarity 31.4%; Pred. No. 15;  
 Matches 27; Conservative 14; Mismatches 25; Indels 20; Gaps 5;

QY 43 QYSAQRORPHQGVILCSAGSTVYAP-----FTGMIYGOEKPYONKNAINNGVRI 93  
 DB 500 QYNAQRSGQMPQ-----AAQAGYQPVLSGQGFQGLI--GVQCPQSQNVINNOGTVY 552  
 QY 94 SGRGFCVYKMFYIKPIKYGPIKKGKEL--GTLPLQKVPYGIQS----HVHIE 140  
 DB 553 QS---VWVSYPTMSSYQVPMTOGSO 574

## RESULT 10

Q9AKK4 PRELIMINARY; PRT; 284 AA.  
 AC 09AKK4;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical 32.4 kDa protein.  
 OS Rickettsia montana.  
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
 OC Rickettsiaceae; Rickettsia.  
 OX NCBI\_TaxID=33991;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21219194; PubMed=11319266;  
 RA Andersson J.O., Andersson S.G.E.;  
 RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes."  
 RL Mol. Biol. Evol. 18:829-839(2001).  
 DR EMBL; AJ293330; CAC33652.1; -.  
 DR InterPro; IPR002886; Peptidase\_M37.  
 DR Pfam; PF01551; Peptidase\_M37.1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 284 AA; 32423 MW; 6708E9F3A54EF1F5 CRC64;

Query Match 9.4%; Score 75.5; DB 2; Length 284;  
 Best Local Similarity 23.6%; Pred. No. 6.5;  
 Matches 26; Conservative 20; Mismatches 35; Indels 29; Gaps 6;

QY 44 YSAQRORPHQGVILCSAGSTVYAPFTGMI--VGQEKPYONKNAINNGVRIISGRGFCV 100  
 DB 171 YKKRRKKCFHSGIDIDQAKKAAPITAAAGVIAARAPDIGNVEIKH-----GKRFY 225



QY 101 KMYIKPIKYK--GPIKKGKGLTLLPLQKVPICQ-----SHWHE 140  
 DB 226 KYAHKEMSVKGNKIKRQO-----FIGIOGSGNATGEHLHFE 264

## RESULT 11

Q9IED5 PRELIMINARY; PRT; 538 AA.  
 AC Q9IED5;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE Env polypeptide (Fragment).  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_Taxid=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YBF51;  
 RA Roques P., Robertson D., Damond F., Souquiere S., Mauciere P.,  
 RT "HIV-1 group O phylogenetic analysis of C2-gp41 region."  
 DR EMBL: AJ133072; CAB96233.1;  
 DR InterPro: IPR000328; Env\_GP41.  
 DR Pfam: PF00516; GP120.1.  
 DR Pfam: PF00517; GP41.1.  
 KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
 FT NON\_TER 1  
 FT SEQUENCE 538 AA; 60777 MW; B3C9E66A233FEFD CRC64;

Query Match 9.4%; Score 75.5; DB 15; Length 538;  
 Best Local Similarity 26.9%; Pred. No. 14;  
 Matches 36; Conservative 13; Mismatches 42; Indels 43; Gaps 6;

QY 25 CAGKSSN--EIRTCRHGGGQYSAORSORPHOGVDILCSAGSTVYAP-----FT 71  
 DB 175 CEGSGCNFKENRT-----NNNTQIPCKLKQYVRSMKSGSLYAPIRGNLCRSMT 227  
 QY 72 GMIVGQEKPYONKNAINGVRISG-----RGFCVMFYIKPIKYKPI--- 114  
 DB 228 GMILQDLKPW-NRNDTNTFRPIGGMKMDIWTELLRKVRIKPFVAPTRIRAPVIGS 286  
 QY 115 -----KKGKGLT 123  
 DB 287 GTREKRAVGLGML 300

## RESULT 12

Q9HS60 PRELIMINARY; PRT; 240 AA.  
 AC Q9HS60;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DE Vng0391C.  
 GN VNG0391C.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae; Halobacterium.  
 NCBI\_Taxid=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20504483; PubMed=11016950;  
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Shroana J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
 RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudich J.L., Jung K.-H.,

RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Edhardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL: AE004996; AAC18948.1;  
 DR InterPro: IPR002761; DDF71.  
 DR Pfam: PF01902; DDF71.1.  
 DR TIGRfams: TIGR00290; Mj0570\_dom; 1.  
 KW Complete proteome.

QY 9 LAGLSTLACGPWANICAKSSNERT-----CDRRGCGYSARSGRPHGVILCSA 62  
 DB 98 VADTLDGAG--VIAAVESDFQTSRMEGMDRLGCELYAPLMQPPRELAAMLA 153

Query Match 9.3%; Score 75; DB 17; Length 240;  
 Best Local Similarity 23.4%; Pred. No. 6.1;  
 Matches 37; Conservative 23; Mismatches 56; Indels 42; Gaps 8;

QY 63 GSTVYAPFTGMIV-----GOEKPYQKN-----AINN--GVRISSGFCVMEYI 105  
 DB 154 G-----FEIRITQVAAYGLDKSWLGRRLDRDALAALALNDEYGVHMLGEGEETLVT 207  
 QY 106 KPIKYKPIKKGKGLTLLPLQKVPICQSHVH-ENC 142  
 DB 208 DAPHMDRPIR-----LDYETVWAGDRGHIEETAC 237

## RESULT 13

Q8RD77 PRELIMINARY; PRT; 270 AA.  
 AC Q8RD77;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DE Membrane proteins related to metalloendopeptidases.  
 GN NLPD OR TTE0169.  
 OS Thermococcuszoochloae tengcongensis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Clostridia;  
 OC Thermococcobacteriales; Thermococcobacteriaceae; Thermococcobacter.  
 NCBI\_Taxid=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4T / JCM11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of T. tengcongensis genome."  
 RL Genome Res. 12:689-700(2002).  
 DR EMBL: AE012990; AAM23470.1;  
 KW Complete proteome.

QY 53 HQGVDIILCSAGSTVYAPFTGMIV-----VGOEKPYQKNALNNGVRISGRGCVMEYIK----- 106  
 DB 168 HKGIDILACKLEPVVAASDGIYSVKYKDPRLGNTVYIKNGI-----WEMVYASLDGN 219

QY 107 -PIKYKPIKKGKGLTLLPLQKVPICQSHVH-----ENCSDSPYAT 150  
 DB 220 IKVKEGDKITKCGQIGEVGDTAKFELAEAGHLELRENGVPIDPTPY 267

## RESULT 14

Q9SUP5 PRELIMINARY; PRT; 444 AA.  
 AC Q9SUP5;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:14 ; Search time 59 Seconds  
(without alignments)  
341.031 Million cell updates/sec

Title: us-09-868-953a-2

Perfect score: 807  
Sequence: 1 MFSTKALLAGLISTALGP.....GIQSHVHINCDSPPAYL 151

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	807	100.0	151	19	AAW52000
2	802	99.4	151	19	AAW23840
3	802	99.4	151	21	AAAB1932
4	796	98.6	151	17	AAK99271
5	793	91.6	681	22	ABG14781
6	702	87.0	151	19	AAW57630
7	645	79.9	133	17	AAK97815
8	642	79.6	133	14	AAK32177
9	430	53.3	326	23	AAE16952
10	294	36.4	34	22	ABB38145

11	294	36.4	54	22	ABB23353
12	294	36.4	54	22	AAW58781
13	294	36.4	54	22	AAW71294
14	294	36.4	54	22	AAI18972
15	294	36.4	54	22	AAW31573
16	294	36.4	54	23	ABG41094
17	256	31.7	54	17	AAK99266
18	196	24.3	41	17	AAK97816
19	132	16.4	27	17	AAK97829
20	95	11.8	20	17	AAK99270
21	94	11.6	18	17	AAK97819
22	90	11.2	19	17	AAK97818
23	89	11.0	19	17	AAK97827
24	86	10.7	18	17	AAK97826
25	82	10.2	25	23	AAE16951
26	77.5	9.6	580	22	AAK38996
27	76	9.4	438	22	ABG02090
28	75	9.3	309	21	AAK29833
29	75	9.3	349	21	AAK29832
30	75	9.3	444	21	AAK29831
31	75	9.3	444	23	ABK93077
32	74.5	9.2	263	23	AAK88022
33	74	9.2	22	23	AAE16949
34	74	9.2	702	7	AAE02230
35	73	9.0	14	17	AAK97825
36	73	9.0	16	23	AAE16950
37	73	9.0	481	19	AAW83216
38	73	9.0	481	22	AAE04758
39	72.5	9.0	168	22	AAK91668
40	72.5	9.0	355	21	AAK27935
41	72.5	9.0	370	20	AAK27934
42	72	8.9	159	22	AAK43029
43	70.5	8.7	625	19	AAW83200
44	70.5	8.7	625	19	AAW69958
45	70.5	8.7	625	19	AAW68294

ALIGNMENTS

RESULT 1	
AAW52000	AAW52000 standard; Protein; 151 AA.
ID	
XX	
AC	AAW52000;
XX	
DT	15-MAR-2002 (first entry)
XX	
DE	Human Mim-1 homologous protein.
XX	
KW	Human; Mim-1 homologue.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Misc-difference 1
FT	Location/Qualifiers
FT	/note= "encoded by ATO"
FT	Misc-difference 58
FT	/note= "encoded by ATO"
XX	
XX	KR97065715-A.
XX	
PD	13-OCT-1997.
XX	
PF	29-MAR-1996; 96KR-0009150.
XX	
PR	29-MAR-1996; 96KR-0009150.
XX	
PA	(KOAD ) KOREA INST SCI & TECHNOLOGY.
XX	
PI	Choi IS, Chung TH, Kim JH, Lee YH, Lee IA, Choi YG, Song JC;
XX	
XX	Lim JS;
XX	

Protein #5352 enco  
Human brain expres  
Human bone marrow  
Peptide #5406 enco  
Peptide #5610 enco  
Human peptide enco  
Leukocyte cell-der  
Bovine chondromodu  
Leukocyte cell-der  
Bovine chondromodu  
Bovine chondromodu  
Bovine chondromodu  
Chicken myb induce  
Human polypeptide  
Novel human diagno  
Arabidopsis thalia  
Arabidopsis thalia  
Herbicidally activ  
Mouse IIR protein.  
Chicken myb induce  
Dihydroxyacetone-S  
Bovine chondromodu  
Chicken myb induce  
Human h-RAB-R. Ho  
Human vesicle traf  
C glutamincum prote  
Arabidopsis thalia  
Arabidopsis thalia  
Propionibacterium  
Murine osteoclast  
Murine NF-kB recep  
Murine NF-kB recep

DR WPI: 1998-492656/42.  
 DR N-PSDB: ABA03485.  
 XX  
 PT Mim-1 homologous gene derived from human tissue NoAbstract -  
 XX  
 PS Claim 1: Page 2; 4pp; Korean.  
 CC The present invention relates to the human Mim-1 homologous gene. The  
 CC present sequence is the Mim-1 protein.  
 XX  
 SQ Sequence 151 AA:  
 Query Match 100.0%; Score 807; DB 19; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 2, 2e-83;  
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGQGYSAORSQRPBGVDILC 60  
 DB 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGQGYSAORSQRPBGVDILC 60  
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120  
 DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120  
 QY 121 GTLLPLQKVPYPGIOSHVHIENCSDSDPTAYL 151  
 DB 121 GTLLPLQKVPYPGIOSHVHIENCSDSDPTAYL 151

RESULT 2  
 AAW23840  
 ID AAW23840 standard; Protein; 151 AA.

AAW23840;

22-MAY-1998 (first entry)

Human LECT2.

Human; leukocyte-derived chemotaxin 2; LECT2; monoclonal antibody;

Immunossay; diagnosis; hepatitis; liver cirrhosis.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 58 /label= Val, Ile

WO9745451-A1.

04-DEC-1997.

26-MAY-1997; 97MO-JP01775.

27-MAY-1996; 96JP-0132160.

(MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.

Arai T;

WPI: 1998-032582/03.

N-PSDB: AAT76922.

Monoclonal antibodies which recognise human leukocyte-derived  
 Chemotaxin 2 - for diagnostic assay of this chemotaxin in hepatitis  
 and other diseases

Claim 1; Pages 20-21; 43pp; Japanese.

The present sequence is human leukocyte-derived chemotaxin 2

(LECT2).

Novel monoclonal antibodies recognising human LECT2 are produced by  
 culture of hybridoma G2A5D7 (FERM P-15638), hybridoma A1G1C6

CC (FERM P-15639), hybridoma 5C5 (FERM P-15640), hybridoma H12D10D6  
 CC (FERM P-15641) and/or hybridoma 89F2 (FERM P-16229) obtained by  
 CC fusion of spleen cells from mice immunised with human LECT2 as  
 CC antigen, with mouse myeloma cells. The antibodies are used as the  
 CC 1st (immobilised) antibody in ELISA immunoassay for human LECT2,  
 CC for the diagnosis of diseases such as hepatitis and liver  
 CC cirrhosis.  
 XX  
 SQ Sequence 151 AA:

Query Match 99.4%; Score 802; DB 19; Length 151;

Best Local Similarity 99.3%; Pred. No. 8e-83;

Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGQGYSAORSQRPBGVDILC 60

DB 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGQGYSAORSQRPBGVDILC 60

QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120

DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVYKMFYIKPIYKGPICKGKEL 120

QY 121 GTLLPLQKVPYPGIOSHVHIENCSDSDPTAYL 151

DB 121 GTLLPLQKVPYPGIOSHVHIENCSDSDPTAYL 151

RESULT 3

AAB11932  
 ID AAB11932 standard; Protein; 151 AA.

AAB11932;

20-NOV-2000 (first entry)

Human leukocyte-derived chemotaxin 2 (LECT2).

Human leukocyte-derived chemotaxin 2; LECT2; bone resorption inhibitor;

leukocyte activating factor; hypercalcaemia; osteoporosis; bone fracture;

lumbago; lumbodysnia; hyperparathyroidism; Paget's disease.

Homo sapiens.

Key Location/Qualifiers

Protein 19..151 /note= "Mature human LECT2"

Misc-difference 56 /label= Val, Ile

WO200037093-A1.

29-JUN-2000.

20-DEC-1999; 99MO-JP07152.

22-DEC-1998; 98JP-0363727.

(NAIN-) NAT INST INFECTIOUS DISEASES.

(SUZU/) SUZUKI K.

Suzuki K, Yamagoe S, Yamakawa T;

WPI: 2000-442530/38.

Leukocyte activating protein factor LECT2 or its derivative with bone  
 resorption inhibitory activity, useful for efficacious therapy of  
 hypercalcaemia, osteoporosis, bone fracture, lumbodysnia,  
 hyperparathyroidism and Paget's disease -

Claim 2; Page 14-15; 20pp; Japanese.

This sequence represents human leukocyte-derived chemotaxin 2 (LECT2), a  
 leukocyte activating factor. The invention relates to the use of LECT2



CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

CC Sequence 681 AA:

Query Match 91.6%; Score 739; DB 22; Length 681;  
 Best Local Similarity 100.0%; Pred. No. 8.7e-75;  
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 ALAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILCSAGSTVYAPFTGMTV 75  
 DB 546 ALAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILCSAGSTVYAPFTGMTV 605  
 QY 76 GQEKPYQNKNAINGVRISGRFCVGMFYIKPIYKGPICKGKGLTLLPLQKYPGIGS 135  
 DB 606 GQEKPYQNKNAINGVRISGRFCVGMFYIKPIYKGPICKGKGLTLLPLQKYPGIGS 665  
 QY 136 HVHIENCSSDPTAYL 151  
 DB 666 HVHIENCSSDPTAYL 681

RESULT 6  
 AAM57630  
 ID AAM57630 standard; Protein; 151 AA.

XX AC AAM57630;  
 XX DT 27-AUG-1998 (first entry)  
 XX DE Chondromodulin-II protien.  
 XX KW Cow; chondromodulin-II; Chm-II; osteoclast growth; cell growth factor;  
 XX OS cartilage growth; bone fracture healing.  
 XX OS Bos taurus.

FX Key Location/Qualifiers  
 FT Peptide 1..18  
 FT Protein /note= "signal peptide"  
 FT /note= "9..151"  
 FT /note= "mature chondromodulin-II"

PN JP10146189-A.

XX PD 02-JUN-1998.

XX PF 15-NOV-1996; 96JP-0304942.

XX PR 15-NOV-1996; 96JP-0304942.

XX PA (MITU ) MITSUBISHI CHEM CORP.

XX DR WPI; 1998-369864/32.

XX DR N-PSDB; AAV24457.

XX PT DNA coding chondromodulin-II protein - stimulates increase of  
 XX cartilage, and hence healing of bone fractures

XX PS Claim 2; Page 14; 17pp; Japanese.

XX CC This sequence is the bovine chondromodulin-II (Chm-II) protein of  
 CC the invention, which has the following properties: (1) it is

CC water-soluble, and comprises a polypeptide with a molecular weight of  
 CC 16 kD by SDS-PAGE; (2) it has an activity of growing osteoclast alone or  
 CC in the co-presence of another cell growth factor; (3) it has an activity  
 CC of promoting differentiating function in cartilage cell; and (4) it has  
 CC an activity of growing cartilage cell alone or in the co-presence of  
 CC another cell growth factor. The protein stimulates the increase of  
 CC cartilage, and hence the healing of bone fractures. The Chm-II can supply  
 CC recombinant Chm-II stably.

XX Sequence 151 AA:

Query Match 87.0%; Score 702; DB 19; Length 151;  
 Best Local Similarity 86.1%; Pred. No. 1.8e-71;  
 Matches 130; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

QY 1 MFSTKALLAGLSTLAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILC 60  
 DB 1 MFSTKALLAGLSTLAGPWANICAGKSSNEIRTCDRHGGCGOYSAGRSQRPBGVDILC 60  
 QY 61 SAGSTVYAPFTGMTVGOEKPYQNKNAINGVRISGRFCVGMFYIKPIYKGPICKGKGL 120  
 DB 61 SAGSTVYAPFTGMTVGOEKPYQNKNAINGVRISGRFCVGMFYIKPIYKGPICKGKGL 120  
 QY 121 GTLLPLQKYPGIGSHVHIENCSSDPTAYL 151  
 DB 121 GTLLPLQKYPGIGSHVHIENCSSDPTAYL 151

RESULT 7  
 AAR97815  
 ID AAR97815 standard; protein; 133 AA.

XX AC AAR97815;  
 XX DT 22-AUG-1996 (first entry)  
 XX DE Bovine chondromodulin-II.  
 XX KW bone metabolism; osteoclast formation; bone resorption; disorder;  
 XX KW disease; treatment; chondrocyte growth factor; chondromodulin;  
 XX KW fibroblast growth factor.  
 XX OS Bos taurus.

PN JP08027020-A.

XX PD 30-JAN-1996.

XX PF 19-JUL-1994; 94JP-0166917.

XX PR 19-JUL-1994; 94JP-0166917.

XX PA (MITU ) MITSUBISHI CHEM CORP.

XX DR WPI; 1996-173892/18.

XX PD Agent for promoting osteoclast formation and bone absorption  
 XX PF containing chondrocyte growth factor, e.g. chondromodulin, as active  
 XX PT component

XX PS Claim 4; Page 6; 6pp; Japanese.

XX CC Bovine chondromodulin-II has mol. wt. 16 kD and promotes chondrocyte  
 XX growth (opt. in the presence of fibroblast growth factor). The  
 XX CC protein is the preferred chondrocyte growth factor for use as active  
 XX ingredient in a novel agent for promoting osteoclast formation and  
 XX bone resorption. Such an agent is intended for treating metabolic  
 XX bone diseases.

XX Sequence 133 AA:

Query Match 79.9%; Score 645; DB 17; Length 133;  
 Best Local Similarity 87.2%; Pred. No. 4.4e-65;

	Matches	116;	Conservative	8;	Mismatches	9;	Indels	0;	Gaps	0;
QY	19	GPAPNACAGKSSN	EIRCTDRHCGGY	SAQRSGRPHQGV	DILCSAGSTVYA	AFPTGMIVG	OE	78		
Db	1	GPWAIICAGKSSN	EIRCTDGHCGGY	TQAQRNQKLHG	VDVLCSDGSTVYA	AFPTGKIMGE		60		
QY	79	KPQPNKNAINNGV	ISGSGFCVKKMFY	IKPIKTKGPIK	KGEKLGTLPLQ	KVYPCGIQSHV		138		
Db	61	KPKPNKNAINNGV	ISGGFCIKMFY	IKPIKTKGSIK	KEKLGTLPLQ	KVYPCGIQSHH		120		
QY	139	IENCDSDDPTAYL		151						
Db	121	IENCDSLDPITYL		133						

RESULT 8	
AAR32177	
ID	AAR32177 standard; protein; 133 AA.
XX	
AC	AAR32177;
XX	
DT	04-JUN-1993 (first entry)
XX	
DE	Human foetal cartilage chondromodulin II.
XX	
KW	stimulating chondrocyte growth; fractures; treatment; cartilage; disorders.
OS	Homo sapiens.
XX	
PN	EP526883-A.
XX	
PD	10-FEB-1993.
XX	
PF	05-AUG-1992; 92EP-0113342.
XX	
PR	05-AUG-1991; 91JP-0195495.
XX	
PA	(MITU ) MITSUBISHI KASEI CORP.
XX	
PI	Hiraki Y, Kamizono A, Kondo J, Suzuki F, Tanaka H;
PI	Teranishi Y;
DR	WPI; 1993-046832/06.
PT	New chondromodulin-II protein - stimulates chondrocyte growth, used for treating fractures and cartilage disorders
PS	Claim 1; Page 11; 18pp; English.
XX	
CC	This sequence represents human foetal cartilage chondromodulin II. It has a mol wt. of 16KD by SDS-PAGE, the ability to stimulate chondrocytes in the presence or absence of RGF and the ability to promote differential potency of chondrocytes. The protein was fractionated and purified on the basis of the uptake of thymidine into cartilage cells. The sequence of the full protein was obtained by aligning N-terminal sequences and internal sequences. Tests revealed that the protein possessed a potent stimulating effect on chondrocyte growth ie. uptake of radioactive thymidine in presence of 1200ng/ml chondromodulin was increased 3.5 fold of that observed in the absence of RGF and 1.4 fold of that observed in the presence of 0.4ng RGF.
Sequence	133 AA;
130	

[illegible]

OY	79	KRYONKNA.INNNGVRIASRGCCVMKFYIKP.IKYRGP.IKKGEKLGT.LLPLOKVAPG.IOSHHV	138
Dd	61	KRYKNMNA.INNNGVRISGGRCIMKFYIKPIKYKS.IKKGEKLGT.LLPLOKVAPG.IOSHHT	120
OY	139	IENCSSSDPTAYL	151
Dd	121	IENCDSLDPPTYL	133

XX	AEI16952 standard; Protein; 326 AA.
XX	AEI16952;
DT	18-APR-2002 (first entry)
DE	Chicken myb induced myeloid protein (mim-1).
KW	Chicken: myb induced myeloid protein; mim-1; bone formation; osteoblast;
KM	osteoclast-secreted chemokine-like protein; bone resorption; osteoclast;
KW	bone degradation; osteoporosis; metastatic bone disease; osteopathic;
KM	protein therapy; cytostatic.
OS	Gallus gallus.
XX	
PN	WO200197833-A1.
PD	27-DEC-2001.
PF	19-JUN-2001; 2001WO-US19575.
PR	19-JUN-2000; 2000US-212271P.
PA	(UABR-) UAB RES FOUND.
PI	Williams JP, McDonald JM, McKenna MA;
DR	WPI; 2002-098195/13.
PT	Regulating osteoblast function (e.g. bone resorption activity), useful
PT	for regulating bone mass or structural integrity, or treating e.g.
PT	osteoporosis, by stimulating secretion of mim-1 protein from the
PS	osteoclasts -
XX	
XX	Claim 2; Fig 3B; 49pp; English.
CC	The invention relates to the fields of biochemical endocrinology and
CC	regulation of bone formation and degradation. More specifically, the
CC	invention relates to the regulation of osteoblast function by the
CC	osteoclast-secreted chemokine-like protein myb induced myeloid protein
CC	(mim-1). Mim-1 protein is used for regulating osteoblast function,
CC	inducing bone resorption activity of osteoclasts, inducing new bone
CC	synthesis, or regulating recruitment and/or differentiation of
CC	osteoblast precursor cells to provide an essential mechanism for
CC	regulating the mass and structural integrity of bone. Mim-1 protein
CC	is particularly useful for treating a variety of pathological bone
CC	diseases, e.g. osteoporosis and metastatic bone diseases. Mim-1
CC	protein is used in protein therapy. The present sequence is chicken
CC	mim-1 protein.
XX	
XX	Sequence 326 AA;

```
Query Match      53.3% Score 430; DB 23; Length 326;  
Best Local Similarity 50.3%; Pred. No. 3.7e-40;  
Matches    80; Conservative   26; Mismatches   39; Indels    14; Gaps     2.  
  
QY          7 LLLAGLSTALGP-----WANTCAGKSSNEIRTCDRHGCQYSQRSOR--P 52  
              : : : : : | : : : : | : : : : | : : : : |  
Db          7 IALLSTVFARQWEVHPQQQGHHMQICSGNPFNRRCDCRRGCSNGYASRGGKEE 66  
              : : : : : | : : : : | : : : : | : : : : |  
QY          53 HGVVDLCISGSIIYAFTFGNIYGQEKPYOKNNAINNCVRISRGECVCYKMFIKDPTKK 112  
              |:|:|:|:|:|:|:|:|:|:|:|:|:|:~::~::|:|:|:|:|:|:|:|:|:|:|:|:|
```

DB 67 HKGVNICTDSIYVAFWFGQLSGPIRFHNGNAIDDOVQISGSGYCVKLCIHPRIYHG 126  
 OY 113 PIKKGKLGTLPLQKVYVPGIOSHVHENCSSPTAYL 151  
 DB 127 QIQKGQQLGRMLPMQKVFPGIVSHIHVENCDSPTAYL 165

RESULT 10  
 ABB38145  
 ID ABB38145 standard; Peptide; 54 AA.  
 AC ABB38145;  
 DT 04-FEB-2002 (first entry)  
 DE Peptide #5651 encoded by human foetal liver single exon probe.  
 XX Human; foetal liver; gene expression; single exon nucleic acid probe.  
 XX Homo sapiens.  
 XX W0200157277-A2.  
 XX 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US00669.  
 PE 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI: 2001-483447/52.  
 DR Human genome-derived single exon nucleic acid probes useful for  
 XX analyzing gene expression in human fetal liver -  
 XX  
 PS Claim 27; SEQ ID NO 30780; 639pp + sequence listing; English.  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC measuring and displaying gene expression in samples derived from human  
 CC fetal liver. The present sequence is a peptide encoded by a single exon  
 CC nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 54 AA;  
 SO

Query Match 36.4%; Score 294; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVKMFYIRPIKKGKLGTLPLQKVYVPGIOSHVHENCSSPTAYL 151  
 DB 1 FCVKMFYIRPIKKGKLGTLPLQKVYVPGIOSHVHENCSSPTAYL 54

DT 23-JAN-2002 (first entry)  
 DE Protein #5352 encoded by probe for measuring heart cell gene expression.  
 XX  
 XX Human; gene expression; heart; microarray; vascular system;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease.  
 XX  
 XX Homo sapiens.  
 XX W0200157274-A2.  
 XX 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US00666.  
 PE 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI: 2001-488899/53.  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 XX hearts -  
 XX  
 PS Claim 15; SEQ ID NO 25123; 530pp; English.  
 CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart (see  
 CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
 CC probe. The probes may be used for predicting, measuring and displaying  
 CC gene expression in samples derived from the human heart via microarrays.  
 CC By measuring gene expression, the probes are useful for predicting,  
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
 CC human heart and vascular system e.g. cardiovascular disease,  
 CC hypertension, cardiac arrhythmias and congenital heart disease.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 54 AA;  
 SO

Query Match 36.4%; Score 294; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVKMFYIRPIKKGKLGTLPLQKVYVPGIOSHVHENCSSPTAYL 151  
 DB 1 FCVKMFYIRPIKKGKLGTLPLQKVYVPGIOSHVHENCSSPTAYL 54

RESULT 12  
 AAM58781  
 ID AAM58781 standard; Protein; 54 AA.  
 AC AAM58781;  
 DT 05-NOV-2001 (first entry)  
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30886.  
 XX  
 XX Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer.  
 XX



OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 FI  
 DR WPI; 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 XX  
 PS Example 4; SEQ ID NO: 30886; 650pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 CC  
 SQ Sequence 54 AA:  
 Query Match 36.4%; Score 294; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 98 FCVKMFYIKPIKYKGIKGEKLTLLPQKYVPGIQSHVHIENCSSDPAYL 151  
 DB 1 FCVKMFYIKPIKYKGIKGEKLTLLPQKYVPGIQSHVHIENCSSDPAYL 54  
 RESULT 13  
 AAM71294  
 ID AAM71294 standard; Protein: 54 AA.  
 XX  
 AC AAM71294;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31600.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 OS  
 PA WO200157276-A2.  
 XX  
 PI  
 DR  
 PF 09-AUG-2001.  
 XX  
 PR 30-JAN-2001; 2001WO-US00668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.

PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 FI  
 DR WPI; 2001-488900/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 XX  
 PS Example 4; SEQ ID NO: 31600; 658pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 CC  
 SQ Sequence 54 AA:  
 Query Match 36.4%; Score 294; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 98 FCVKMFYIKPIKYKGIKGEKLTLLPQKYVPGIQSHVHIENCSSDPAYL 151  
 DB 1 FCVKMFYIKPIKYKGIKGEKLTLLPQKYVPGIQSHVHIENCSSDPAYL 54  
 RESULT 14  
 AAM18972  
 ID AAM18972 standard; Protein: 54 AA.  
 XX  
 AC AAM18972;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #5406 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe: human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 OS  
 PA WO200157278-A2.  
 XX  
 PI  
 DR  
 PF 09-AUG-2001.  
 XX  
 PR 30-JAN-2001; 2001WO-US00670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 FI  
 DR WPI; 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 27; SEQ ID NO 23798; 487pp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SENP: see AAI10068-AAI28459). The present sequence is a peptide encoded

CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX

SQ Sequence 54 AA:

Query Match 36.4%; Score 294; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 8.6e-26;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 151  
 DB 1 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 54

RESULT 15

AAM31573

ID AAM31573 standard; Protein; 54 AA.

AC AAM31573;

DT 17-OCT-2001 (first entry)

DE Peptide #5610 encoded by probe for measuring placental gene expression.

KW Probe: microarray; human; placenta; antenatal diagnosis;

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 31842; 654bp; English.

XX

SQ Sequence 54 AA:

Query Match 36.4%; Score 294; DB 22; Length 54;

Best Local Similarity 100.0%; Pred. No. 8.6e-26;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 151  
 DB 1 FCVAMFYIKPIKYGPIKKGKGLGTLPLQKYVPGIOSHVHIENCSSDPTAYL 54

Search completed: March 25, 2003, 17:19:37  
 Job time : 60 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:14 ; Search time 59 Seconds  
(without alignments)  
341.031 Million cell updates/sec

Title: US-09-868-953A-1  
Perfect score: 804  
Sequence: 1 MFSTKALLAGLISTALGP.....GIQSHVHENCSDPTAYL 151

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	802	99.8	151	19	AAAM52000
2	802	99.8	151	19	AAAM23840
3	802	99.8	151	21	AAAB11932
4	792	98.5	151	17	AAAP9271
5	734	91.3	681	22	ABG14781
6	698	86.8	151	19	AAAM57630
7	641	79.7	133	17	AAAR7815
8	638	79.4	133	14	AAAR32177
9	426	53.0	326	23	AAAL6952
10	294	36.6	54	22	ABAB38145

11	294	36.6	54	22	ABAB23353
12	294	36.6	54	22	AAAM58781
13	294	36.6	54	22	AAAM71294
14	294	36.6	54	22	AAAM18972
15	294	36.6	54	22	AAAM31573
16	294	36.6	54	23	ABG41094
17	252	31.3	54	17	AAAR9266
18	192	23.9	41	17	AAAR97816
19	132	16.4	27	17	AAAR97816
20	95	11.8	20	17	AAAR9270
21	94	11.7	18	17	AAAR97819
22	89	11.1	19	17	AAAR97827
23	86	10.7	18	17	AAAR97826
24	86	10.7	19	17	AAAR97818
25	78	9.7	25	23	AAE16951
26	76	9.5	438	22	ABG02090
27	75.5	9.4	580	22	AAAM38996
28	74	9.2	22	23	AAE16949
29	74	9.2	702	7	AAAP60230
30	73	9.1	14	17	AAAR97825
31	73	9.1	16	23	AAE16950
32	73	9.1	159	22	AAU43029
33	72.5	9.0	715	20	AAAY05625
34	71.5	8.9	706	23	ABAB92784
35	70.5	8.8	642	23	ABAB92784
36	70.5	8.8	680	22	AAAB92854
37	70.5	8.8	680	23	ABAB92854
38	70	8.7	309	21	AAAG20833
39	70	8.7	349	21	AAAG20833
40	70	8.7	444	21	AAAG20831
41	69.5	8.7	444	23	ABAB93077
42	69.5	8.6	168	22	AAAG91968
43	69.5	8.6	263	23	AAU88022
44	69	8.6	14	17	AAAR97824
45	68	8.5	214	22	AAU41669

## ALIGNMENTS

RESULT 1	AAAM52000	standard; Protein; 151 AA.
ID	AAAM52000;	
AC	AAAM52000;	
XX		
DT	15-MAR-2002	(first entry)
DE	Human Mim-1 homologous protein.	
XX		
KW	Human: Mim-1 homologue.	
XX		
OS	Homo sapiens.	
XX		
FT	key	Location/Qualifiers
FT	Misc-difference 1	/note= "encoded by APO"
FT	Misc-difference 58	/note= "encoded by APO"
FT	Misc-difference 58	/note= "encoded by APO"
XX		
PN	KR97065715-A.	
XX		
PD	13-OCT-1997.	
XX		
PF	29-MAR-1996;	96KR-0009150.
XX		
PR	29-MAR-1996;	96KR-0009150.
XX		
PA	(KOAD ) KOREA INST SCI & TECHNOLOGY.	
XX		
Chol IS;	Chung TH, Kim JH, Lee YH, Lee IA, Choi YG, Song JC;	
PI	Lim JS;	
XX		

Protein #5352 enco  
Human brain expres  
Human bone marrow  
Peptide #5406 enco  
Peptide #5610 enco  
Human peptide enco  
Leukocyte cell-der  
Bovine chondromodu  
Bovine chondromodu  
Leukocyte cell-der  
Bovine chondromodu  
Bovine chondromodu  
Bovine chondromodu  
Chicken myb induce  
Novel human diagno  
Human polypeptide  
Chicken myb induce  
Dihydroxyacetone-s  
Bovine chondromodu  
Chicken myb induce  
Protonibacterium  
HIV-1 group O Isol  
Herbically activ  
Novel human protei  
Human protein sequ  
Novel human protei  
Arabidopsis thalia  
Arabidopsis thalia  
Herbically activ  
C glutamicum prote  
Mouse Irf8 protein.  
Bovine chondromodu  
Protonibacterium

DR WPI: 1998-492656/42.  
 DR N-PSDB: ABA03485.  
 XX  
 XX Mim-1 homologous gene derived from human tissue NCabstract -  
 PS Claim 1; Page 2; 4pp; Korean.  
 CC The present invention relates to the human Mim-1 homologous gene. The  
 CC present sequence is the Mim-1 protein.  
 XX  
 SQ Sequence 151 AA;  
 Query Match 99.8%; Score 802; DB 19; Length 151;  
 Best Local Similarity 99.3%; Pred. No. 2.5e-87;  
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MFSTALLAGLISLALAGPWANICAGKSSNEIRTCRHGGCGQYSAQSRPHGVDXLC 60  
 DB 1 MFSTALLAGLISLALAGPWANICAGKSSNEIRTCRHGGCGQYSAQSRPHGVDXLC 60  
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYGPIKKGKEL 120  
 DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYGPIKKGKEL 120  
 QY 121 GTLLPLQKYVPGIQSHVHIENCDSDEPTAYL 151  
 DB 121 GTLLPLQKYVPGIQSHVHIENCDSDEPTAYL 151  
 RESULT 2  
 AAW23840  
 ID AAW23840 standard; Protein: 151 AA.  
 XX  
 AC AAW23840;  
 XX  
 DT 22-MAY-1998 (first entry)  
 DE Human LECT2.  
 XX  
 KW Human; leukocyte-derived chemotaxin 2; LECT2; monoclonal antibody;  
 KW immunoassay; diagnosis; hepatitis; liver cirrhosis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT MISC-difference 58  
 FT /label= Val, Ile  
 XX  
 FN WO9745451-A1.  
 XX  
 PD 04-DEC-1997.  
 XX  
 PE 26-MAY-1997; 97WO-JP01775.  
 XX  
 PR 27-MAY-1996; 96UP-0132160.  
 XX  
 PA (MEDI-) MEDICAL & BIOLOGICAL LAB CO LTD.  
 XX  
 PI Arai T;  
 XX  
 DR WPI: 1998-032582/03.  
 DR N-PSDB: AAT76922.  
 XX  
 PT Monoclonal antibodies which recognise human leukocyte-derived  
 PT chemotaxin 2 - for diagnostic assay of this chemotaxin in hepatitis  
 PT and other diseases  
 XX  
 PS Claim 1; Pages 20-21; 43pp; Japanese.  
 XX  
 CC The present sequence is human leukocyte-derived chemotaxin 2  
 CC (LECT2).  
 CC Novel monoclonal antibodies recognising human LECT2 are produced by  
 CC culture of hybridoma G2A5D7 (FERM P-15638), hybridoma A1G1C6

CC (FERM P-15639), hybridoma 5C5 (FERM P-15640), hybridoma H12D10D6  
 CC (FERM P-15641) and/or hybridoma 89F2 (FERM P-16229) obtained by  
 CC fusion of spleen cells from mice immunised with human LECT2 as  
 CC antigen, with mouse myeloma cells. The antibodies are used as the  
 CC 1st (immobilised) antibody in ELISA immunoassay for human LECT2,  
 CC for the diagnosis of diseases such as hepatitis and liver  
 CC cirrhosis.  
 XX  
 SQ Sequence 151 AA;  
 Query Match 99.8%; Score 802; DB 19; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-87;  
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MFSTALLAGLISLALAGPWANICAGKSSNEIRTCRHGGCGQYSAQSRPHGVDXLC 60  
 DB 1 MFSTALLAGLISLALAGPWANICAGKSSNEIRTCRHGGCGQYSAQSRPHGVDXLC 60  
 QY 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYGPIKKGKEL 120  
 DB 61 SAGSTVYAPFTGMIVGOEKPYQNKNAINGVRISGRFCVAMFYIKPIKYGPIKKGKEL 120  
 QY 121 GTLLPLQKYVPGIQSHVHIENCDSDEPTAYL 151  
 DB 121 GTLLPLQKYVPGIQSHVHIENCDSDEPTAYL 151  
 RESULT 3  
 AAB11932  
 ID AAB11932 standard; Protein: 151 AA.  
 XX  
 AC AAB11932;  
 XX  
 DT 20-NOV-2000 (first entry)  
 DE Human leukocyte-derived chemotaxin 2 (LECT2).  
 XX  
 KW Human leukocyte-derived chemotaxin 2; LECT2; bone resorption inhibitor;  
 KW leukocyte activating factor; hypercalcaemia; osteoporosis; bone fracture;  
 KW lumbago; lumbodysia; hyperparathyroidism; Paget's disease.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 19..151  
 FT /note= "Mature human LECT2"  
 FT MISC-difference 56  
 FT /label= Val, Ile  
 XX  
 FN WO200037093-A1.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PE 20-DEC-1999; 99WO-JP07152.  
 XX  
 PR 22-DEC-1998; 98UP-0363727.  
 XX  
 PA (NAIN-) NAT INST INFECTIOUS DISEASES.  
 XX  
 PI (SUZU/) SUZUKI K.  
 XX  
 DR WPI: 2000-442530/38.  
 DR  
 XX  
 PT Leukocyte activating protein factor LECT2 or its derivative with bone  
 PT resorption inhibitory activity, useful for efficacious therapy of  
 PT hypercalcaemia, osteoporosis, bone fracture, lumbodysia,  
 PT hyperparathyroidism and Paget's disease -  
 XX  
 PS Claim 2; Page 14-15; 20pp; Japanese.  
 XX  
 CC This sequence represents human leukocyte-derived chemotaxin 2 (LECT2), a  
 CC leukocyte activating factor. The invention relates to the use of LECT2

CC and derivatives thereof as bone resorption inhibitors. The invention also  
 CC encompasses a method of screening candidate bone resorption inhibitors  
 CC using LECT2, and the inhibitors thus identified. Human LECT2 was isolated  
 CC from undifferentiated bone cells, and its ability to inhibit osteoclast-  
 CC mediated bone resorption was examined via a pit assay. Human LECT2 caused  
 CC 50% inhibition of bone resorption at 1 microgram/ml and 100% inhibition  
 CC at 10 micrograms/ml. LECT2 or its derivatives may be used for the  
 CC treatment of hypercalcaemia, osteoporosis, bone fractures, lumbago  
 CC (lumbodysnia), hyperparathyroidism and Paget's disease.

XX Sequence 151 AA:

Query Match 99.8%; Score 802; DB 21; Length 151;  
 Best Local Similarity 100.0%; Pred. No. 2.5e-87;  
 Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAQRORPHQGVXLC 60  
 |||||||  
 Db 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAQRORPHQGVXLC 60  
 QY 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKNFYIKPIYKGPPIKKEKL 120  
 |||||||  
 Db 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKNFYIKPIYKGPPIKKEKL 120  
 QY 121 GTLLPLQKYYPGIGSHVHIENCSSDPTAYL 151  
 |||||||  
 Db 121 GTLLPLQKYYPGIGSHVHIENCSSDPTAYL 151

#### RESULT 4

AAR99271 ID AAR99271 standard; Protein; 151 AA.

XX AAR99271:  
 DT 13-DEC-1996 (first entry)  
 XX

DE Leukocyte cell-derived chemotaxin LECT2b.

XX LECT2b; LECT2a; leukocyte cell-derived chemotaxin; cytokine;  
 KW neutrophil; chemotaxis; cancer; therapy; diagnosis.

XX Homo sapiens.

OS Key Location/Qualifiers

PH MISC-difference 42

FT /note= "Residue 42 may also be ile, owing to a  
 FT polymorphism in the nucleotide sequence"

PN EP723016-A2.

XX 24-JUL-1996.

XX 27-NOV-1995; 95EP-0402670.

XX 28-NOV-1994; 94JP-0293233.

XX (SUZU/) SUZUKI K.

XX Arita S, Mizuno S, Suzuki K, Yamakawa Y, Yamagoe S;

XX WPI: 1996-335477/34.

XX N-PSDB; AAT35260.

PT New human leukocyte cell-derived chemotaxin cpds. - used for the  
 PT detection and therapy of cancer and disorders of the cytokine  
 PT network

PS Claim 2; Page 13; 17pp; English.

XX Novel human leukocyte cell-derived chemotaxin LECT2b (AAR99271) has  
 CC neutrophil chemotactic activity. Its amino acid sequence was  
 CC deduced from a cDNA clone (AAT35260) derived from

CC phytohaemagglutinin-activated T-cell leukaemia cell line SKW-3.  
 CC Recombinant LECT2a can be produced, pref. as a fusion protein, in  
 CC transformed host cells, esp. E. coli, yeast, insect cells, CHO,  
 CC CV-1, 293, C127, 3T3, L-929, HeLa and SKW-3 cells. LECT2b and  
 CC LECT2a (see also AAR9266-70) are useful for the diagnosis, therapy  
 CC and prediction of cancer and disorders of the cytokine network.

XX Sequence 151 AA:

Query Match 98.5%; Score 792; DB 17; Length 151;  
 Best Local Similarity 98.0%; Pred. No. 4e-86;  
 Matches 148; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAQRORPHQGVXLC 60  
 |||||||  
 Db 1 MFSTALLAGLSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAQRORPHQGVXLC 60  
 QY 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKNFYIKPIYKGPPIKKEKL 120  
 |||||||  
 Db 61 SAGSTVYAPFTGMIVGEEKPYQNKNAIINGVRIISGRGFCVKNFYIKPIYKGPPIKKEKL 120  
 QY 121 GTLLPLQKYYPGIGSHVHIENCSSDPTAYL 151  
 |||||||  
 Db 121 GTLLPLQKYYPGIGSHVHIENCSSDPTAYL 151

#### RESULT 5

ABG14781 ID ABG14781 standard; Protein; 681 AA.

XX ABG14781:

DT 18-FEB-2002 (first entry)  
 XX

DE Novel human diagnostic protein #14772.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001MO-US08631.

XX 31-MAR-2000; 2000US-0540217.

XX 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YF.

XX WPI: 2001-639362/73.

XX N-PSDB; AAS78968.

PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity

PS Claim 20; SEQ ID NO 45140; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SO Sequence 681 AA:  
 Query Match 91.3%; Score 734; DB 22; Length 681;  
 Best Local Similarity 99.3%; Pred. No. 2.5e-78;  
 Matches 135; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 16 ALAGPMANICAGKSSNEIRFCDRHGGCGYSQRSORPHQGVYDLCASGTVYAPFTGMTV 75  
 DB 546 ALAGPMANICAGKSSNEIRFCDRHGGCGYSQRSORPHQGVYDLCASGTVYAPFTGMTV 605  
 QY 76 GOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKLTLLPLQKVYPGIGS 135  
 DB 606 GOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKLTLLPLQKVYPGIGS 665  
 QY 136 HVHIENCSSDPTAYL 151  
 DB 666 HVHIENCSSDPTAYL 681  
 RESULT 6  
 AAM57630  
 ID AAM57630 standard; Protein: 151 AA.  
 AC AAM57630;  
 XX 27-AUG-1998 (first entry)  
 DT Chondromodulin-II protein.  
 DE Chondromodulin-II protein.  
 XX Cow; chondromodulin-II; ChM-II; osteoclast growth; cell growth factor;  
 KW cartilage growth; bone fracture healing.  
 XX Bos taurus.  
 OS  
 XX Key Location/Qualifiers  
 FH Peptide 1..18  
 FT /note= "signal peptide"  
 FT 9..151  
 FT Protein /note= "mature chondromodulin-II"  
 PN JP10146189-A.  
 XX 02-JUN-1998.  
 PD  
 XX 15-NOV-1996; 96JP-0304942.  
 PE  
 XX 15-NOV-1996; 96JP-0304942.  
 PR  
 XX (MITU) MITSUBISHI CHEM CORP.  
 PA  
 XX WPI: 1998-369864/32.  
 DR N-PSDB; AAV244457.  
 XX DNA coding chondromodulin-II protein - stimulates increase of  
 PT cartilage, and hence healing of bone fractures  
 XX  
 PS Claim 2: Page 14; 17pp: Japanese.  
 XX This sequence is the bovine chondromodulin-II (ChM-II) protein of  
 CC the invention, which has the following properties: (1) It is

CC water-soluble, and comprises a polypeptide with a molecular weight of  
 CC 16 kD by SDS-PAGE; (2) it has an activity of growing osteoclast alone or  
 CC in the co-presence of another cell growth factor; (3) it has an activity  
 CC of promoting differentiating function in cartilage cell; and (4) it has  
 CC an activity of growing cartilage cell alone or in the co-presence of  
 CC another cell growth factor. The protein stimulates the increase of  
 CC cartilage, and hence the healing of bone fractures. The ChM-II can supply  
 CC recombinant ChM-II stably.  
 CC  
 SO Sequence 151 AA:  
 Query Match 86.8%; Score 698; DB 19; Length 151;  
 Best Local Similarity 86.1%; Pred. No. 6.3e-75;  
 Matches 130; Conservative 7; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 MESTRALDLAAGLSTALAGPMANICAGKSSNEIRFCDRHGGCGYSQRSORPHQGVYDLC 60  
 DB 1 MESTGTLTLLAALISPALAGPMAICAGKSSNEIRFCDRHGGCGYTAORQKHLHGVYDLC 60  
 QY 61 SAGSTYVAPFTGMIVGOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKL 120  
 DB 61 SAGSTYVAPFTGMIVGOEKPYONKNAIINGVRIISGRGFCVKMFYIKPIKYPKIKKGEKL 120  
 QY 121 GTLLPLQKVYPGIGSHVHIENCSSDPTAYL 151  
 DB 121 GTLLPLQKVYPGIGSHVHIENCSSDPTAYL 151  
 RESULT 7  
 AAR97815  
 ID AAR97815 standard; protein: 133 AA.  
 AC AAR97815;  
 XX 22-AUG-1996 (first entry)  
 DT Bovine chondromodulin-II.  
 DE Bovine chondromodulin-II.  
 XX bone metabolism; osteoclast formation; bone resorption; disorder;  
 KW disease; treatment; chondrocyte growth factor; chondromodulin;  
 KW fibroblast growth factor.  
 XX Bos taurus.  
 OS  
 XX JP08027020-A.  
 PN 30-JAN-1996.  
 PD  
 XX 19-JUL-1994; 94JP-0166917.  
 PF  
 XX 19-JUL-1994; 94JP-0166917.  
 PR  
 XX (MITU) MITSUBISHI CHEM CORP.  
 PA  
 XX WPI: 1996-173892/18.  
 DR  
 XX Agent for promoting osteoclast formation and bone absorption  
 PT containing chondrocyte growth factor, e.g. chondromodulin, as active  
 PT component  
 XX  
 PS Claim 4: Page 6; 6pp: Japanese.  
 XX Bovine chondromodulin-II has mol. wt. 16 kD and promotes chondrocyte  
 CC growth (opt. in the presence of fibroblast growth factor). The  
 CC protein is the preferred chondrocyte growth factor for use as active  
 CC ingredient in a novel agent for promoting osteoclast formation and  
 CC bone resorption. Such an agent is intended for treating metabolic  
 CC bone diseases.  
 CC  
 SO Sequence 133 AA:  
 Query Match 79.7%; Score 641; DB 17; Length 133;  
 Best Local Similarity 87.2%; Pred. No. 3.3e-68;







OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains -  
 XX  
 PS Example 4; SEQ ID NO: 30886; 650pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is a protein encoded by one of  
 CC the probes of the invention.  
 XX  
 SQ Sequence 54 AA:  
 Query Match 36.6%; Score 294; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-27;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 98 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIHENCSSDPAYL 151  
 DB 1 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIHENCSSDPAYL 54  
 RESULT 13  
 ID AAM71294  
 ID AAM71294 standard; Protein; 54 AA.  
 XX  
 AC AAM71294;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31600.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488900/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human bone marrow -  
 XX  
 PS Example 4; SEQ ID NO: 31600; 658pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a  
 CC protein encoded by one of the probes of the invention.  
 XX  
 SQ Sequence 54 AA:  
 Query Match 36.6%; Score 294; DB 22; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-27;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 98 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIHENCSSDPAYL 151  
 DB 1 FCVKMFYIKPIKYKGPICKGKGLTLLPLQKVPGIOSHVIHENCSSDPAYL 54  
 RESULT 14  
 ID AAM18972  
 ID AAM18972 standard; Protein; 54 AA.  
 XX  
 AC AAM18972;  
 XX  
 DT 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #5406 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 27; SEQ ID NO 23798; 487pp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SNP: see AAI10068-AA128459). The present sequence is a peptide encoded

CC by one such probe. The SENPs are derived from human Hela cells. The SENPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 54 AA;

Query Match 36.6%; Score 294; DB 22; Length 54;

Best Local Similarity 100.0%; Pred. No. 2.2e-27;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVKMFYIKPIKYGPIKKGKGLTLPLOKVYPGIGSHVHIENCSSDPTAYL 151  
 DB 1 FCVKMFYIKPIKYGPIKKGKGLTLPLOKVYPGIGSHVHIENCSSDPTAYL 54

# RESULT 15

AAM31573  
 ID AAM31573 standard; Protein; 54 AA.

XX AAM31573;

DT 17-OCT-2001 (first entry)

DE Peptide #5610 encoded by probe for measuring placental gene expression.

KW Probe: microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.

OS Homo sapiens.

PN WO200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00663.

PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -

PS Claim 27; SEQ ID No 31842; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP:  
 CC see AAI31315-AAI57546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.

XX Sequence 54 AA;

Query Match 36.6%; Score 294; DB 22; Length 54;

Best Local Similarity 100.0%; Pred. No. 2.2e-27;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 98 FCVKMFYIKPIKYGPIKKGKGLTLPLOKVYPGIGSHVHIENCSSDPTAYL 151  
 DB 1 FCVKMFYIKPIKYGPIKKGKGLTLPLOKVYPGIGSHVHIENCSSDPTAYL 54

Search completed: March 25, 2003, 17:19:36  
 Job time : 62 secs

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:17:45 ; Search time 19 Seconds

(without alignments)  
233.835 Million cell updates/sec

Title: US-09-868-953A-1

Perfect score: 804  
Sequence: 1 MFSTKALLLAGLISTALGP.....GIQSHVHINCDSDDPYAYL 151

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: Issued\_Patents\_AA.\*
- 2: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*
- 3: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*
- 4: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*
- 5: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*
- 6: /cgn2\_6/ptodata/1/1aa/BACKFILES1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	802	99.8	151	2	US-08-563-148E-6
2	802	99.8	151	4	US-09-194-139-1
3	641	79.7	133	1	US-07-924-753-15
4	417	51.9	98	4	US-09-194-139-9
5	252	31.3	54	2	US-08-563-148E-1
6	192	23.9	41	1	US-07-924-753-1
7	132	16.4	27	1	US-07-924-753-14
8	95	11.8	20	2	US-08-563-148E-5
9	94	11.7	18	1	US-07-924-753-4
10	89	11.1	19	1	US-07-924-753-12
11	86	10.7	18	1	US-07-924-753-11
12	86	10.7	19	1	US-07-924-753-3
13	74	9.2	702	6	5240838-15
14	73	9.1	14	1	US-07-924-753-10
15	69	8.6	14	1	US-07-924-753-9
16	67.5	8.4	625	3	US-08-996-139-15
17	67.5	8.4	625	4	US-08-995-659-15
18	67.5	8.4	625	4	US-09-215-649A-15
19	67.5	8.4	625	4	US-09-577-780-15
20	67	8.3	10182	4	US-09-134-001C-3159
21	66.5	8.3	2285	4	US-09-308-375-2
22	65.5	8.1	1241	4	US-09-040-774-2
23	64	8.0	428	3	US-09-118-319-5
24	63.5	7.9	2595	4	US-09-036-987A-2
25	63.5	7.9	2595	4	US-09-370-700-2
26	63	7.8	14	4	US-09-194-139-10
27	63	7.8	2004	1	US-08-375-709-15

28	63	7.8	2004	1	US-08-752-929-15	Sequence 15, Appl
29	63	7.8	2004	4	US-09-090-793-9	Sequence 9, Appl
30	62.5	7.8	406	1	US-08-293-778-24	Sequence 24, Appl
31	62.5	7.8	406	1	US-08-295-411-5	Sequence 5, Appl
32	62.5	7.8	406	2	US-08-955-471-5	Sequence 5, Appl
33	62.5	7.8	406	5	PCT-US92-10242-5	Sequence 5, Appl
34	62.5	7.8	433	3	US-08-926-842B-61	Sequence 61, Appl
35	62.5	7.8	444	2	US-08-475-845-2	Sequence 2, Appl
36	62.5	7.8	444	2	US-08-327-690-2	Sequence 2, Appl
37	62.5	7.8	444	2	US-08-660-289-2	Sequence 2, Appl
38	62.5	7.8	444	2	US-08-537-807-2	Sequence 2, Appl
39	62.5	7.8	444	2	US-08-871-003-2	Sequence 2, Appl
40	62.5	7.8	444	3	US-08-464-233-2	Sequence 2, Appl
41	62.5	7.8	444	4	US-09-189-607-2	Sequence 2, Appl
42	62.5	7.8	444	4	US-09-378-907-2	Sequence 2, Appl
43	62.5	7.8	444	5	PCT-US94-05779-2	Sequence 2, Appl
44	62.5	7.8	466	1	US-07-882-202A-4	Sequence 4, Appl
45	62.5	7.8	466	1	US-08-021-615A-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
US-08-563-148E-6  
Sequence 6, Application US/08563148E  
Patent No. 5929224  
GENERAL INFORMATION:  
APPLICANT: Kazuo SUZUKI et al.  
TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSSEE: 663-2, Shilgi  
STREET: Misaki-machi  
CITY: Isumi-gun  
STATE: Chiba-ken  
COUNTRY: JAPAN  
ZIP: 259-45  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS v.5  
SOFTWARE: Word Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/563,148E  
FILING DATE: No. 5929224ember 27, 1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 293233/1994  
FILING DATE: 28-NOV-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: C. Bruce Hamburg  
REGISTRATION NUMBER: 22,389  
REFERENCE/DOCKET NUMBER: F-5230  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)966-2340  
FAX: (212)953-7733  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: single strand  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: intermediate fragment  
ORIGINAL SOURCE:  
ORGANISM: human  
TISSUE TYPE: T-cell derived leukemia cells  
PUBLICATION INFORMATION:  
AUTHORS: Kazuo SUZUKI et al.  
TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
RELEVANT RESIDUES IN SEQ ID NO: 6: FROM 1 to 151  
US-08-563-148E-6

Query Match 99.8%; Score 802; DB 2; Length 151;  
Best Local Similarity 99.3%; Pred. No. 1,1e-89;  
Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQVDXLC 60  
DB 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQVDXLC 60  
QY 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKPKIKGK 120  
DB 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKPKIKGK 120  
QY 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151  
DB 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151

## RESULT 2

US-09-194-139-1  
Sequence 1, Application US/09194139A  
Patent No. 6306608  
GENERAL INFORMATION:  
APPLICANT: Arai, Takao  
TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,  
TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME  
FILE REFERENCE: US application 09/194,139  
CURRENT APPLICATION NUMBER: US/09/194,139A  
CURRENT FILING DATE: 1998-11-25  
EARLIER APPLICATION NUMBER: JP 8-132160  
EARLIER FILING DATE: 1996-05-27  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 151  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: residue 58 is either Valine or Isoleucine  
US-09-194-139-1

Query Match 99.8%; Score 802; DB 4; Length 151;  
Best Local Similarity 100.0%; Pred. No. 1,1e-89;  
Matches 151; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQVDXLC 60  
DB 1 MESTKALLAGLISALAGWANICAGKSSNEIRTCDRHGGCGQYSAQRSPRHQVDXLC 60  
QY 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKPKIKGK 120  
DB 61 SAGSTVYAPFTGMIVGQEKPYQNKAINNGVIRISGRGFCVKMFYIKPIKPKIKGK 120  
QY 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151  
DB 121 GTLLPLQKYYPGIQSHVHIENCSSDPTAYL 151

## RESULT 3

US-07-924-753-15  
Sequence 15, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 133 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE:  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:

CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:

US-07-924-753-15

Query Match 79.7%; Score 641; DB 1; Length 133;  
Best Local Similarity 87.2%; Pred. No. 3.2e-70;  
Matches 116; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 19 GPWANCAGKSSNEIRTCDRHGGCGQYSAQRSPRHQVDXLC SAGSTVYAPFTGMIVGQ 78  
DB 1 GPWANCAGKSSNEIRTCDRHGGCGQYSAQRSPRHQVDXLC SAGSTVYAPFTGMIVGQ 60

QY 79 KPYONKNAIINGVRIISGRGFCVAKMEFYIKPIKYKGPPIKKEKLTLLPLQKYVPGIOSHVH 138  
 Db 61 KPYONKNAIINGVRIISGRGFCVAKMEFYIKPIKYKGIKKEKLTLLPLQKYVPGIOSHVH 120  
 QY 139 IENCDDSDPTAVL 151  
 Db 121 IENCDDSDPTAVL 133

RESULT 4

US-09-194-139-9  
 ; Sequence 9, Application US/09194139A  
 ; Patent No. 6306608  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Arai, Takao  
 ; TITLE OF INVENTION: ANTI-HUMAN LECT2 ANTIBODY, CELLS PRODUCING THE SAME,  
 ; TITLE OF INVENTION: AND METHOD AND KIT FOR ASSAYING THE SAME  
 ; FILE REFERENCE: US Application 09/194,139  
 ; CURRENT APPLICATION NUMBER: US/09/194,139A  
 ; EARLIER FILING DATE: 1998-11-25  
 ; EARLIER FILING DATE: 1996-05-27  
 ; NUMBER OF SEQ ID NOS: 10  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 9  
 ; LENGTH: 98  
 ; TYPE: PRT  
 ; ORGANISM: Bovine  
 US-09-194-139-9

Query Match 51.9%; Score 417; DB 4; Length 98;  
 Best Local Similarity 64.6%; Pred. No. 3.5e-43;  
 Matches 84; Conservative 6; Mismatches 8; Indels 32; Gaps 4;

QY 19 GPMANICAGKSSNEIRTCRHHGCGGYTAORNRKHOGVDVLCASGTYAPFTG 78  
 Db 1 GPMANICAGKSSNEIRTCRHHGCGGYTAORNRKHOGVDVLCASGTYAPFTG-IMGOE 59  
 QY 79 KPYONKNAIINGVRIISGRGFCVAKMEFYIKPIKYKGPPIKKEKLTLLPLQKYVPGIOSHVH 138  
 Db 60 KPYONKNAIINGVRIISGRGFCVAKMEFYIKPIKYKGIKKEKLTLLPLQKYVPGIOSHVH 120  
 QY 139 IENCDDSDPT 148  
 Db 89 IENCDDSDPT 98

RESULT 5

US-08-563-148E-1  
 ; Sequence 1, Application US/08563148E  
 ; Patent No. 5929224  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kazuo SUZUKI et al.  
 ; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
 ; NUMBER OF SEQUENCES: 12  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: 663-2, Shligi  
 ; STREET: Misaki-machi  
 ; CITY: Isumi-gun  
 ; STATE: Chiba-ken  
 ; COUNTRY: JAPAN  
 ; ZIP: 299-45  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: MS-DOS V.5  
 ; SOFTWARE: Word Perfect 5.1  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/563,148E  
 ; FILING DATE: No. 5929224ember 27, 1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 293233/1994  
 ; FILING DATE: 28-NOV-1994

ATTORNEY/AGENT INFORMATION:  
 ; NAME: C. Bruce Hamburg  
 ; REGISTRATION NUMBER: 22,389  
 ; REFERENCE/DOCKET NUMBER: F-5230  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212)986-2340  
 ; TELEFAX: (212)953-7733  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 54 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single strand  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; FRAGMENT TYPE: N-terminal fragment  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: human  
 ; TISSUE TYPE: T-cell derived leukemia cells  
 ; PUBLICATION INFORMATION:  
 ; AUTHORS: Kazuo SUZUKI et al.  
 ; TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES  
 ; RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 to 54  
 US-08-563-148E-1

Query Match 31.3%; Score 252; DB 2; Length 54;  
 Best Local Similarity 83.3%; Pred. No. 1.7e-23;  
 Matches 45; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 19 GPMANICAGKSSNEIRTCRHHGCGGYTAORNRKHOGVDVLCASGTYAPFTG 72  
 Db 1 GPMANICAGKSSNEIRTCRHHGCGGYTAORNRKHOGVDVLCASGTYAPFTG 54

RESULT 6

US-07-924-753-1  
 ; Sequence 1, Application US/07924753  
 ; Patent No. 5270303  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fujio SUZUKI et al.  
 ; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Wenderoth, Lind & Ponack  
 ; STREET: 805 Fifteenth Street, N.W., #700  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: MS-DOS  
 ; SOFTWARE: DisplayWrite  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/924,753  
 ; FILING DATE: 19920804  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Warren M. Cheek, Jr.  
 ; REGISTRATION NUMBER: 33,367  
 ; REFERENCE/DOCKET NUMBER:  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-8850  
 ; TELEFAX:  
 ; TELEPHONE:  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 41 amino acids  
 ; TYPE: AMINO ACID  
 ; STRANDEDNESS: single

TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: N-terminal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-1

Query Match 23.9%; Score 192; DB 1; Length 41;  
Best Local Similarity 82.9%; Pred. No. 2.3e-16;  
Matches 34; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 19 GPMANICAGSSNEIRTCDRHGGGQYSAORSORPHQGVYDL 59  
|||||  
Db 1 GPMATICAGKSSNEIRTCDSHGCGQYTAQRNOKLHGQYDV 41

RESULT 7  
US-07-924-753-14  
Sequence 14, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804

CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 27 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-14

Query Match 16.4%; Score 132; DB 1; Length 27;  
Best Local Similarity 92.6%; Pred. No. 2.6e-09;  
Matches 25; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 111 KPIKKGKLTLLPLQKYYPGIQSHV 137  
|||  
Db 1 KGIKKGKLTLLPLQKYYPGIQSHI 27

RESULT 8  
US-08-563-148E-5  
Sequence 5, Application US/08563148E  
Patent No. 5929224  
GENERAL INFORMATION:  
APPLICANT: Kazuo SUZUKI et al.

```

; TITLE OF INVENTION: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES: 12
; ADDRESSEE: 663-2, Shitagi
; STREET: Misaki-machi
; CITY: Isumi-gun
; STATE: Chiba-ken
; COUNTRY: JAPAN
; ZIP: 299-45
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS v. 5
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/563,148E
; FILING DATE: NO. 5929224ember 27, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 293233/1994
; FILING DATE: 28-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: C. Bruce Hamburg
; REGISTRATION NUMBER: 22,389
; REFERENCE/DOCKET NUMBER: F-5230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)986-2340
; TELEFAX: (212)953-7733
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single strand
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: intermediate fragment
; ORIGINAL SOURCE:
; ORGANISM: human
; TISSUE TYPE: T-cell derived leukemia cells
; PUBLICATION INFORMATION:
; AUTHORS: Kazuo SUZUKI et al.
; TITLE: NOVEL ACTIVATING FACTOR OF LEUKOCYTES
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 to 20
; US-08-563-148E-5

Query Match 11.8%; Score 95; DB 2; Length 20;
Best Local Similarity 85.0%; Pred. No. 5.4e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 129 VYPGIQSHVHENCDSPT 148
Db 1 VYPGIQSHIHENCXLDSPDT 20

RESULT 9
US-07-924-753-4
; Sequence 4, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; City: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TEXT:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-924-753-4

Query Match 11.7%; Score 94; DB 1; Length 18;
Best Local Similarity 88.9%; Pred. No. 6.2e-05;
Matches 16; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 129 VYPGIQSHVHENCDSPT 146
Db 1 VYPGIQSHIHENCXLDSD 18

RESULT 10
US-07-924-753-12
; Sequence 12, Application US/07924753

```

Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:  
UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:

FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-12  
Query Match 11.1%; Score 89; DB 1; Length 19;  
Best Local Similarity 84.2%; Pred. No. 0.00027;  
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 68 APTGMIVGQEKPKYKNKA 86  
DB 1 APTGKIMGQEKPKYKNKA 19  
RESULT 11  
US-07-924-753-11  
Sequence 11, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: Fujio SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:



```

; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-924-753-11

Query Match
Best Local Similarity 10.7%; Score 86; DB 1; Length 18;
Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 79 KPYONKNAINNGVRISG 95
DB 1 KPYONKNAINNGVRISG 17

RESULT 12
; US-07-924-753-3
; Sequence 3, Application US/07924753
; Patent No. 5270303
; GENERAL INFORMATION:
; APPLICANT: Fujio SUZUKI et al.
; TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Displaywrite
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,753
; FILING DATE: 19920804
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

; HYPOTHETICAL:
; ANTI-SENSE:
; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-924-753-3

Query Match
Best Local Similarity 10.7%; Score 86; DB 1; Length 19;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 53 HGVVDVLCDSAGSTVYAPF 70
DB 2 HGVVDVLCDSAGSTVYAPF 19

RESULT 13
; 5240838-15
; Patent No. 5240838
; APPLICANT: LEDEROER, ADRIANUS M.; MAAT, JAN; VERRIPS, CORNELIS
; T.; VISSER, CHRISTIAN; JANOWICZ, ZBIGNIEW A.; HOLLENBERG, CORNELIS P.
; TITLE OF INVENTION: REGULATORY SEQUENCES OF ALCOHOL OXIDASE
; (MOX) AND DIHYDROXYACETONESYNTASE (DAS) OF HANSENULA POLYMORPHA
; NUMBER OF SEQUENCES: 19
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 300,211
; FILING DATE: 23-JAN-1989
; APPLICATION NUMBER: 759,315
; FILING DATE: 26-JUL-1985
; SEQ ID NO:15;
; LENGTH: 702
; 5240838-15

Query Match
Best Local Similarity 9.2%; Score 74; DB 6; Length 702;
Matches 33; Conservative 22; Mismatches 48; Indels 34; Gaps 7;

QY 11 GLISTALAGPWANICAGKSSNEIRTCDRHGGGOYSAQRSDRPHQGVDXLCSAGSTVYAPF 70

```

Db 525 GLFOKAVELPFSSILS-LSRNEVL-----CYLASRAQRNRNAGYILDEAENAEVQI 575  
QY 71 TGMTVGOEKPYNKNAINNVRISGRF-----CYKMFYIKPIKKGP1--KKGEL 120  
Db 576 IG-VGAEMEFPADKA-----KITGRKFRVRVLSIPCTRLPDEOSIGYRSVLRDRGROV 628  
QY 121 GTLLPLQKVPQISQSHV 137  
Db 629 PTVV-----VDAHV 637

RESULT 14  
US-07-924-753-10  
Sequence 10, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: FUJIO SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:  
FRAGMENT TYPE: internal fragment  
ORIGINAL SOURCE:  
ORGANISM: bovine  
STRAIN:  
INDIVIDUAL ISOLATE:  
DEVELOPMENTAL STAGE:  
HAPLOTYPE:  
TISSUE TYPE: fetal cartilage  
CELL TYPE:  
CELL LINE:  
ORGANELLE:  
IMMEDIATE SOURCE:  
LIBRARY:  
CLONE:  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT:  
MAP POSITION:

UNITS:  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS:  
TITLE:  
JOURNAL:  
VOLUME:  
ISSUE:  
PAGES:  
DATE:  
DOCUMENT NUMBER:  
FILING DATE:  
PUBLICATION DATE:  
RELEVANT RESIDUES IN SEQ ID NO:  
US-07-924-753-10

Query Match 9.1%; Score 73; DB 1; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 KLGTLLPLQKVPQ 132  
Db 1 KLGTLLPLQKVPQ 14

RESULT 15  
US-07-924-753-9  
Sequence 9, Application US/07924753  
Patent No. 5270303  
GENERAL INFORMATION:  
APPLICANT: FUJIO SUZUKI et al.  
TITLE OF INVENTION: CHONDROMODULIN-II PROTEIN  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Wenderoth, Lind & Ponack  
STREET: 805 Fifteenth Street, N.W., #700  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 inch, 500 kb  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: DisplayWrite  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,753  
FILING DATE: 19920804  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Warren M. Cheek, Jr.  
REGISTRATION NUMBER: 33,367  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-8850  
TELEFAX:  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 14 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL:  
ANTI-SENSE:

```

; FRAGMENT TYPE: internal fragment
; ORIGINAL SOURCE:
; ORGANISM: bovine
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE: fetal cartilage
; CELL TYPE:
; CELL LINE:
; ORGANELLAE:
; IMMEDIATE SOURCE:
; LIBRARY:
; CLONE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION:
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
US-07-924-753-9

Query Match      8.6%; Score 69; DB 1; Length 14;
Best Local Similarity 92.9%; Pred. No. 0.048;
Matches 13; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 103 FYIKPIKYKSPDK 116
Db 1 FYIKPIKYKSPDK 14

```

Search completed: March 25, 2003, 17:23:00  
Job time : 28 secs



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:21:15 ; Search time 12 Seconds  
(without alignments)  
672.733 Million cell updates/sec

Title: US-09-868-953a-1

Perfect score: 804  
Sequence: 1 MESTKALLAGLSTALGP.....GIQSHVHIENCSDSPATYL 151

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 221153 seqs, 53462247 residues

Total number of hits satisfying chosen parameters: 221153

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PC1\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PC105\_PUBCOMB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	540	67.2	113	US-09-864-570-9	Sequence 9, Appl1
2	540	67.2	113	US-10-157-457-9	Sequence 9, Appl1
3	426	53.0	326	US-09-884-570-8	Sequence 8, Appl1
4	426	53.0	326	US-10-157-457-8	Sequence 8, Appl1
5	348	43.3	115	US-09-884-570-11	Sequence 11, Appl1
6	348	43.3	115	US-10-157-457-11	Sequence 11, Appl1
7	343	42.7	115	US-09-884-570-10	Sequence 10, Appl1
8	343	42.7	115	US-10-157-457-10	Sequence 10, Appl1
9	294	36.6	54	US-09-864-761-38651	Sequence 38651, A
10	78	9.7	25	US-09-884-570-7	Sequence 7, Appl1
11	78	9.7	25	US-10-157-457-7	Sequence 7, Appl1
12	74	9.2	22	US-09-884-570-5	Sequence 5, Appl1
13	74	9.2	22	US-10-157-457-5	Sequence 5, Appl1
14	73	9.1	16	US-09-884-570-6	Sequence 6, Appl1
15	73	9.1	16	US-10-157-457-6	Sequence 6, Appl1
16	69.5	8.6	168	US-09-738-626-5722	Sequence 5722, Ap
17	67.5	8.4	625	US-09-877-650-15	Sequence 15, Appl1
18	67.5	8.4	625	US-10-166-232A-2	Sequence 2, Appl1
19	67.5	8.4	625	US-09-871-856-15	Sequence 15, Appl1

20	66.5	8.3	2285	US-09-932-183A-2	Sequence 2, Appl1
21	65.5	8.1	502	US-09-738-626-4772	Sequence 4722, Ap
22	65.5	8.1	1241	US-09-759-632-2	Sequence 2, Appl1
23	64.5	8.0	197	US-09-924-340-66	Sequence 66, Appl1
24	64.5	8.0	197	US-09-992-600A-66	Sequence 66, Appl1
25	64.5	8.0	197	US-09-731-872-276	Sequence 276, Appl1
26	64	8.0	598	US-09-854-845-41	Sequence 41, Appl1
27	64	8.0	697	US-09-854-845-25	Sequence 25, Appl1
28	63.5	7.9	151	US-09-925-301-874	Sequence 874, Appl1
29	63.5	7.9	615	US-09-768-799A-2	Sequence 2, Appl1
30	63	7.8	332	US-09-925-300-1064	Sequence 1064, Appl1
31	63	7.8	1615	US-09-931-375A-2	Sequence 2, Appl1
32	62.5	7.8	375	US-09-864-761-36336	Sequence 36336, A
33	62.5	7.8	406	US-10-109-498-1	Sequence 1, Appl1
34	62.5	7.8	603	US-09-854-845-39	Sequence 39, Appl1
35	62.5	7.8	691	US-10-101-921-4	Sequence 4, Appl1
36	62.5	7.8	691	US-09-925-731-2	Sequence 2, Appl1
37	62.5	7.8	702	US-09-854-845-23	Sequence 23, Appl1
38	61.5	7.6	723	US-10-133-912-18	Sequence 18, Appl1
39	61.5	7.6	444	US-09-840-795-8	Sequence 8, Appl1
40	61.5	7.6	492	US-10-087-667-6	Sequence 6, Appl1
41	61.5	7.6	655	US-09-978-295A-64	Sequence 64, Appl1
42	61.5	7.6	655	US-09-756-854-2	Sequence 2, Appl1
43	61.5	7.6	655	US-09-978-697-64	Sequence 64, Appl1
44	61.5	7.6	655	US-10-041-574-2	Sequence 2, Appl1
45	61.5	7.6	655	US-09-978-192A-64	Sequence 64, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-884-570-9  
Sequence 9, Application US/09884570  
Patent No. US20020114779A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294  
CURRENT APPLICATION NUMBER: US/09/884,570  
CURRENT FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/212,271  
PRIOR FILING DATE: 2000-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 9  
LENGTH: 113  
TYPE: PRT  
ORGANISM: unknown  
FEATURE:  
LOCATION: 38..165  
OTHER INFORMATION: amino acid sequence of region of Lect2  
OTHER INFORMATION: protein homologous to repeat sequences of N  
US-09-884-570-9  
Query Match  
Best Local Similarity 67.2%; Score 540; DB 10; Length 113;  
Matches 97; Conservative 9; Mismatches 7; Indels 0; Gaps 0;  
QY 39 HGGQYSAQRQPHQGVYDLCASGTYVAPFTGMVGEKPYONKNAINGVRSRGCF 98  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 1 YGGQYSAQRQPHQGVYDLCASGTYVAPFTGMVGEKPYONKNAINDGIRSGCF 60  
QY 99 CVKMFYIKPIKYGKIKGKIGTLPLQKIVPGIOSHVHIENCSDSPATYL 151  
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||  
Db 61 CVKIFIKPIKYGKIKGKIGTLPLQKIVPGIOSHVHIENCSDSPATYL 113  
RESULT 2  
US-10-157-457-9  
Sequence 9, Application US/10157457





PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 38651  
LENGTH: 54  
TYPE: prt  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC004763.1  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.8  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.9  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.8  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.8  
OTHER INFORMATION: SWISSPROT HIT: O14960, EVALU8 8.00e-17  
US-09-864-761-38651

Query Match 36.6%; Score 294; DB 10; Length 54;  
Best Local Similarity 100.0%; Pred. No. 4.9e-27;  
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 98 FCVKMFYIPKIKKPKKGEKLTLLPLQVYVPGIQSHVHENCDSPTAVL 151  
Db 1 FCVKMFYIPKIKKPKKGEKLTLLPLQVYVPGIQSHVHENCDSPTAVL 54  
RESULT 10  
US-09-884-570-7  
Sequence 7, Application US/09884570  
Patent No. US20020114779A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294  
CURRENT APPLICATION NUMBER: US/09/884,570  
CURRENT FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/212,271  
PRIOR FILING DATE: 2000-06-19  
NUMBER OF SEQ ID NOS: 11

SEQ ID NO 7  
LENGTH: 25  
TYPE: prt  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 219..243  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-09-884-570-7

Query Match 9.7%; Score 78; DB 10; Length 25;  
Best Local Similarity 56.0%; Pred. No. 0.017;  
Matches 14; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Oy 55 GVDXLCASGTYVAPFGMIVGOEK 79  
Db 1 GVDXLCASGTYVAPFGMIVGOEK 25

RESULT 11  
US-10-157-457-7  
Sequence 7, Application US/10157457  
Patent No. US20020150568A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294CIP  
CURRENT APPLICATION NUMBER: US/10/157,457  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: US 09/884,570  
PRIOR FILING DATE: 2001-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 7  
LENGTH: 25  
TYPE: prt  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 219..243  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-10-157-457-7

Query Match 9.7%; Score 78; DB 12; Length 25;  
Best Local Similarity 56.0%; Pred. No. 0.017;  
Matches 14; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Oy 55 GVDXLCASGTYVAPFGMIVGOEK 79  
Db 1 GVDXLCASGTYVAPFGMIVGOEK 25

RESULT 12  
US-09-884-570-5  
Sequence 5, Application US/09884570  
Patent No. US20020114779A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294  
CURRENT APPLICATION NUMBER: US/09/884,570  
CURRENT FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/212,271  
PRIOR FILING DATE: 2000-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 5  
LENGTH: 22  
TYPE: prt  
ORGANISM: Gallus gallus  
FEATURE:



LOCATION: 94..115  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-09-884-570-5

Query Match 9.2%; Score 74; DB 10; Length 22;  
Best Local Similarity 61.9%; Pred. No. 0.042;  
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 81 YONKAINNGVRISGRFCVK 101  
DB 2 FHNQNAIDDGVOISGSGYCVK 22

RESULT 13  
US-10-157-457-5  
Sequence 5, Application US/10157457  
Patent No. US20020150568A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294CIP  
CURRENT APPLICATION NUMBER: US/10/157,457  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: US 09/884,570  
PRIOR FILING DATE: 2001-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 5  
LENGTH: 22  
TYPE: PRT  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 94..115  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-10-157-457-5

Query Match 9.2%; Score 74; DB 12; Length 22;  
Best Local Similarity 61.9%; Pred. No. 0.042;  
Matches 13; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

OY 81 YONKAINNGVRISGRFCVK 101  
DB 2 FHNQNAIDDGVOISGSGYCVK 22

RESULT 14  
US-09-884-570-6  
Sequence 6, Application US/09884570  
Patent No. US20020114779A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294  
CURRENT APPLICATION NUMBER: US/09/884,570  
CURRENT FILING DATE: 2001-06-19  
PRIOR APPLICATION NUMBER: US 60/212,271  
PRIOR FILING DATE: 2000-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 6  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 293..308  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-09-884-570-6

Query Match 9.1%; Score 73; DB 10; Length 16;  
Best Local Similarity 73.3%; Pred. No. 0.037;  
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 129 VYPGIQSHVHENC 143  
DB 1 VFPGLIISHVHENC 15

RESULT 15  
US-10-157-457-6  
Sequence 6, Application US/10157457  
Patent No. US20020150568A1  
GENERAL INFORMATION:  
APPLICANT: Williams, John P.  
APPLICANT: McDonald, Jay M.  
APPLICANT: McKenna, Margaret A.  
TITLE OF INVENTION: Osteoclast Secreted Chemokine and Uses Thereof  
FILE REFERENCE: D6294CIP  
CURRENT APPLICATION NUMBER: US/10/157,457  
CURRENT FILING DATE: 2002-05-29  
PRIOR APPLICATION NUMBER: US 09/884,570  
PRIOR FILING DATE: 2001-06-19  
NUMBER OF SEQ ID NOS: 11  
SEQ ID NO 6  
LENGTH: 16  
TYPE: PRT  
ORGANISM: Gallus gallus  
FEATURE:  
LOCATION: 293..308  
OTHER INFORMATION: amino acid sequence tryptic peptide from  
OTHER INFORMATION: mim-1 protein  
US-10-157-457-6

Query Match 9.1%; Score 73; DB 12; Length 16;  
Best Local Similarity 73.3%; Pred. No. 0.037;  
Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 129 VYPGIQSHVHENC 143  
DB 1 VFPGLIISHVHENC 15

Search completed: March 25, 2003, 17:29:44  
Job time : 13 secs



GenCore version 5.1.4\_p5-4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:16:59 ; Search time 28.5 seconds  
(without alignments)  
509.344 Million cell updates/sec

Title: US-09-868-953A-1

Perfect score: 804  
Sequence: 1 MFSTKALLAGLISTALGP.....GIQSHVHENCSDSPFAYL 151

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	64.1	79.7	133	2 JH0270	chondromodulin II
2	42.9	53.4	326	2 A33755	myb-induced myeloid
3	99.5	12.4	305	2 T16567	hypothetical prote
4	80.5	10.0	417	2 C82840	conserved hypothet
5	76.5	9.5	475	2 F64151	hypothetical prote
6	76	9.5	240	2 H84197	hypothetical prote
7	76	9.5	651	2 T42644	hypothetical prote
8	74	9.2	750	2 F90571	lipoprotein (impor
9	73	9.1	350	2 AD0466	hypothetical prote
10	73	9.1	710	1 XJH9FK	formaldehyde trans
11	72.5	9.0	428	2 T36565	probable penicilli
12	71.5	8.9	295	2 AC2451	hypothetical prote
13	71.5	8.9	706	2 T01351	subtilisin-like pr
14	71	8.8	284	2 F97836	probable periplasm
15	70	8.7	444	2 T05614	hypothetical prote
16	69.5	8.6	268	2 D97054	probable membrane
17	69	8.6	1331	1 XORTDH	xanthine dehydroge
18	69	8.6	1434	2 T32647	hypothetical prote
19	68	8.5	184	2 T15360	hypothetical prote
20	68	8.5	310	2 A60967	insulin-like growt
21	68	8.5	762	2 D70407	DNA mismatch repai
22	68	8.5	3010	1 S18030	genome polyprotein
23	67.5	8.4	273	2 E81330	probable periplasm
24	67.5	8.4	312	2 AE2235	hypothetical prote
25	67.5	8.4	355	2 T51414	CONSTANS-like 1 -
26	67.5	8.4	535	2 C36811	hypothetical prote
27	67.5	8.4	535	2 A37994	RPI protein - salm
28	67	8.3	398	2 A35281	integumentary muc
29	67	8.3	554	2 C98133	component of type

30	67	8.3	554	2 AE3154	agrobacterium virtu
31	67	8.3	689	2 G84447	hypothetical prote
32	67	8.3	1052	2 T14343	zinc finger RNA bi
33	67	8.3	1347	2 T45632	helicase homolog F
34	66.5	8.3	2285	2 T12796	probable transglyc
35	66	8.2	320	2 C85440	myb-related protei
36	66	8.2	447	2 B83563	conserved hypothet
37	66	8.2	782	2 S18032	genome polyprotein
38	65.5	8.1	233	2 AD3344	hypothetical prote
39	65.5	8.1	364	2 T12039	cysteine proteinas
40	65.5	8.1	364	2 T46630	cysteine proteinas
41	65.5	8.1	454	2 AE3537	formyl-coenzyme a
42	65.5	8.1	606	2 H97678	hypothetical prote
43	65.5	8.1	606	2 AE2903	GTP-binding lysos
44	65.5	8.1	608	2 AB3562	GTP-binding protei
45	65.5	8.1	621	2 D96554	hypothetical prote

## ALIGNMENTS

```

RESULT 1
JH0270
chondromodulin II - bovine
N:Alternate names: cartilage-derived factor
C:Species: Bos primigenius taurus (cattle)
C>Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 17-Mar-1999
C:Accession: JH0270
R:Hiraki, Y.; Inoue, H.; Kondo, J.; Kamizono, A.; Yoshitake, Y.; Shikunami, C.;
J. Biol. Chem. 271, 22657-22662, 1996
A>Title: A novel growth-promoting factor derived from fetal bovine cartilage, cl
A:Reference number: JH0270; MUID:96394331; PMID:8798437
A:Accession: JH0270
A:Molecule type: protein
A:Residues: 1-133 <HIR>
A:Experimental source: cartilage cell
C:Comment: This protein stimulates proteoglycan synthesis in the growth plate ch

Query Match          79.7%: Score 641; DB 2; Length 133;
Best Local Similarity 87.2%: Pred. No. 6.7e-59;
Matches 116; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

QY 19 GPMANICAGKSSNEIRTCDRHCGGOYSAORPHOGVDXLCASGVYAPFTGIVGOE 78
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 GPMALICAGKSSNEIRTCDCGCGYTAQRNOKLHOGVDVLCSDGSTVYAPFTGIMGOE 60

QY 79 KPYQNKAINNGVRSRGKGFYKMPYIKPIKYGPIKKGKGLPLPLQKYYPGQSHVH 138
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KPYRNKAINNGVRSISGGFCIKMFYIKPIKYGSIKKGKGLPLPLQKYYPGQSHIH 120

QY 139 IENCDSDPFAYL 151
    ||||| ||||| ||
Db 121 IENCDSDPFAYL 133

RESULT 2
A33755
myb-induced myeloid protein 1 (mim-1) precursor - chicken
N:Alternate names: arginine-specific ADP-ribosyltransferase endogenous substrate
C:Species: Gallus gallus (chicken)
C>Date: 30-Mar-1990 #sequence_revision 30-Mar-1990 #text_change 04-Sep-1998
C:Accession: A33755; S29229
R:Ness, S.A.; Marknell, A.; Graf, T.
Cell 59, 1115-1125, 1989
A>Title: The v-myd oncogene product binds to and activates the promyelocyte-spec
A:Reference number: A33755; MUID:90090611; PMID:2688896
A:Accession: A33755
A:Molecule type: mRNA
A:Residues: 1-326 <NBS>
A:Cross-references: GB:M29449; NID:g212341; PID:g212342
R:Yamada, K.; Tsuchiya, M.; Mishima, K.; Shimoyama, M.
FEBS Lett. 311, 203-205, 1992
A>Title: p13, an endogenous target protein for arginine-specific ADP-ribosyltran

```

[illegible]

```

RESULT 3
T16567
hypothetical protein K05F1.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16567
R:Wohlmann, P.
Submitted to the EMBL Data Library, June 1995
A:Description: The sequence of C. elegans cosmids K05F1.
A:Reference number: Z18537
A:Accession: T16567
A:Status: Preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-305 <R08>
A:Cross-references: EMBL:U29377; NID:q868173; PID:q868182; PIDN:AAA68719.1; CESP:K05F1.4
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:K05F1.4
A:Introns: 44/2; 100/1; 164/1; 188/2; 295/1

Query Match          12.4%; Score 99.5; DB 2; Length 305;
Best Local Similarity 23.4%; Pred. No. 0.014;
Matches 29; Conservative 21; Mismatches 37; Indels 27; Gaps 5;

QY 55 GVDXLCASSTVYAPFTGIVGOEPRY---ONKNAINNGVRIISG---RGFCVKMYIYKP 107
      || | : : || | | : : || : : || | : : || : : || : : || : : ||
Db 52 GVDVACHLDEPRITAPRIEGEMKFW-RPYGGKREKSCADQGVRIEGTGMQGYAVNHISYKL 110
      : : : : : | : : : : : : : : : : : : : : : : : : : : : : : :

QY 108 IKYKPRIKKGEKLTLLPL---OKYVVGIGSHV-----HIENC 142
      : : : : : | : : : : : : : : : : : : : : : : : : : : : : : :
Db 111 SFFGGHVYAGDEIGEALNRYCFNDRQNDVPERHYEIRLYKESRLIDPTNHLQNC 164
      : : : : : | : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
C82840
conserved hypothetical protein XF0167 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 03-Jun-2002
C:Accession: C82840
R:anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing
Native 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A8515; NUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below

```

A:Accession: CB2840  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-417 <SIM>  
A:Cross-references: GB:AE003870; GB:AE003849; NID:q9104955; PIDN:AAE82980.1; GSPD  
A:Experimental source: strain 9a5c  
R:Simpon, A.J.G.; Rehnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga  
B:lonos, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Ca  
s-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Perro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Mar  
A:Authors: Martins, E.M.F.; Matsukuma, A.T.; Menck, C.F.M.; Miracca, E.C.; Miyaki  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palme  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.;  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da  
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore,  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0167  
;Superfamily: hypothetical protein H10409

```

34 PCDNRHGCGGYSNORSORPHOGVDPXLCSSVYVAFPTGMI--VGQCKPYONKNKINNGV 91
Db 249 RLSSTFGRLRRHPILGTRMRMHKGVDAATGTPIMAGADARVQFIGQGRGYONVILNH-- 306

QY 92 RISRGECVMMFYIKPIKYPKIKGKEL--GTLPLQKYYPGIQS----HVHIE 140
Db 307 ---GKGY--RTLIGHMSRF--GKIKAGKIKINGSTVIG---YVGWMTGLATGPHLHYE 352

RESULT 5
F64151
hypothetical protein H10409 - Haemophilus influenzae (strain Rd KW20)
C:Species: Haemophilus influenzae
C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 29-Sep-1999
C:Accession: F64151
R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerla
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Wei
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghagen, N.
Science 269, 496-512, 1995
A:Authors: Gheble, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; V
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae R
A:Reference number: A64000; MUID:95350630; PMID:7542800
A:Accession: F64151
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residue type: <1>A>
A:Cross-references: GB:U32724; GB:U42023; NID:g1573378; PID:AAC22068.1; PID:g157
A:Note: best homolog was a hypothetical protein from Escherichia coli
C:Superfamily: hypothetical protein H10409

Query Match 9.5%; Score 76.5; DB 2; Length 475;
Best Local Similarity 28.7%; Pred.No.5.3;
Matches 27; Conservative 15; Mismatches 41; Indels 11; Gaps 4;

QY 51 RPHOGVDXCSASTVYAPFTGMIVGQCKPYONKNKINNGVRISRGECVMMFYIKP--I 108
Db 346 RPHKGVDFSVSGSTPVYARADGV--EKVAYAGAGGRVYMLKHGREYQVYUMLSKSLV 403

QY 109 KYKGRPKGKELGTLPLQKYYPGIQS--HVHIE 140
Db 404 KAGQTVKKGBRIAL-----SGNTGISGTGPHLHYE 432

RESULT 6
1884197

```



C:Species: *Pichia angusta*  
C:Date: 31-Mar-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Apr-2000  
C:Accession: A23009; S30110  
R:Jankowicz, Z. A.; Eckart, M. R.; Drewke, C.; Roggenkamp, R. O.; Hollenberg, C. P.  
Nucleic Acids Res. 13, 3043-3062, 1985  
A:Title: Cloning and characterization of the DAS gene encoding the major methanol assimilation  
A:Reference number: A23009; MUID:85215670; PMID:2987872  
A:Accession: A23009  
A:Molecule type: DNA  
A:Residues: 1-693; 'RLRPEKGA' <JAN>  
A:Cross-references: GB:X02424  
A:Experimental source: ATCC 34438  
A:Note: This sequence has been revised in reference S30110  
R:Hansen, H.; Didion, T.; Thiemann, A.; Veenhuis, M.; Roggenkamp, R.  
Mol. Gen. Genet. 235, 269-278, 1992  
A:Title: Targeting sequences of the two major peroxisomal proteins in the methylotrophic  
A:Reference number: S30110; MUID:93101130; PMID:1465101  
A:Accession: S30110  
A:Molecule type: DNA  
A:Residues: 667-710 <HAN>  
A:Cross-references: EMBL:X02424  
A:Note: this is a revision to the sequence from reference A23009  
C:Comment: This is the major methanol assimilatory enzyme from this methylotrophic organ  
C:Genetics:  
A:Gene: DAS  
C:Superfamily: transketolase; thiamin pyrophosphate-binding domain homology  
C:Keywords: peroxisome; thiamin pyrophosphate; transferase  
F:18-208/Domain: thiamin pyrophosphate-binding domain homology #status atypical <TPB>  
F:708-710/Region: peroxisome/glyoxysome location signal #status atypical

Query Match 9.1%; Score 73; DB 1; Length 710;  
Best Local Similarity 24.1%; Pred. No. 18;  
Matches 33; Conservative 21; Mismatches 49; Indels 34; Gaps 7;

Qy 11 GLISALAGPMNINICGKSSNEIRTCORHCGOYSAORSPHOGVDXLCASGTYAPF 70  
Db 525 GLFOKAVELPFSITLS-LSRNEVL-----QTLASRAQRNRNAGTYLEAENAEVOT 575

Qy 71 TGMINGOEKPYONKNAINNGVRISGRF-----CYKMFYIKPIKYKGP1--KKGEKL 120  
Db 576 IG--VGAEMEFADKAK-----KITGRKFRTRVLSIPCTRLEDSIGVRSVLRKDGROY 628

Qy 121 GTLPLQKVPYPGIQSHV 137  
Db 629 PTVV-----VDGHV 637

RESULT 11  
T36565  
probable penicillin-binding protein - Streptomyces coelicolor (fragment)  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Nov-2000  
C:Accession: T36565  
R:Seeger, K. J.; Harris, D.; James, K. D.; Parkhill, J.; Barrell, B. G.; Rajandream, M. A.  
submitted to the EMBL Data Library, June 1999  
A:Reference number: Z21584  
A:Accession: T36565  
A:Status: preliminary; translated from GB/EMBL/DBDUT  
A:Molecule type: DNA  
A:Residues: 1-428 <SEP>  
A:Cross-references: EMBL:AL079353; PTDN:CA84563.1; GSPDB:GN00070; SCODEB:SCH17.14  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCODEB:SCH17.14  
C:Superfamily: penicillin-binding protein 1B

Query Match 9.0%; Score 72.5; DB 2; Length 428;  
Best Local Similarity 38.9%; Pred. No. 12;  
Matches 21; Conservative 6; Mismatches 22; Indels 5; Gaps 2;

Qy 43 OYASROSPHOGVDXLCASGTYAPFTGMIV--GOEKPY--ONKNAINNGV 91  
Db 342 QESVHQSLKDHVKNKSDVNAASLIVPEGTGRIVAMGOSKPYGYCKNTEINYSV 395

RESULT 12  
AC2451  
hypothetical protein all5163 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp.  
A:Note: Nostoc sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
C:Accession: AC2451  
R:Kaneko, T.; Nakamura, Y.; Molk, C. P.; Kuritz, T.; Sasamoto, S.; Matanabe, A.; I  
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Ta  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacter  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AC2451  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-295 <KUP>  
A:Cross-references: GB:BA000019; PTDN:BAB76862.1; PID:917134301; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all5163

Query Match 8.9%; Score 71.5; DB 2; Length 295;  
Best Local Similarity 29.6%; Pred. No. 11;  
Matches 34; Conservative 14; Mismatches 40; Indels 27; Gaps 7;

Qy 53 HOGVDXLCASGTY--YAPFTGMIVGOEKPYONKNAIN--NGVRISGRGFCVYKMPYIKPIK 109  
Db 190 HSGVDLLALGTPVEALIPGTVEAFKQGYGNLVIIHNC-----GLQSKYAHLDSDIN 243

Qy 110 YK--GPICKKEKGLTL---LPLOKVPYPGIQSHVHIENCDS-----DPIAYL 151  
Db 244 KVGQKVNQGLGTGTGTGQPTAK-----QPHLHFEVRVSSSLGVAENPKDYL 293

RESULT 13  
T01351  
subtilisin-like proteinase homolog F6N15.3 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 20-Sep-1999  
C:Accession: T01351  
R:Ryan, E.; Edwards, J.; Pape, K.  
submitted to the EMBL Data Library, May 1998  
A:Description: The sequence of A. thaliana F6N15.  
A:Reference number: 214297  
A:Accession: T01351  
A:Status: translated from GB/EMBL/DBDUT  
A:Molecule type: DNA  
A:Residues: 1-706 <RYA>  
A:Cross-references: EMBL:AF069299; MUID:93193311; PID:93193320  
A:Experimental source: cultivar Columbia  
C:Genetics:  
A:Map position: 4  
A:Introns: 10/2; 43/1; 91/1; 126/2; 297/3; 330/2; 418/2; 454/3; 526/1  
A:Note: F6N15.3  
C:Superfamily: subtilisin-like proteinase ag12; subtilisin homology

Query Match 8.9%; Score 71.5; DB 2; Length 706;  
Best Local Similarity 25.0%; Pred. No. 26;  
Matches 33; Conservative 21; Mismatches 35; Indels 43; Gaps 8;

Qy 8 LLAG-----LISTALAGFWANICAGKSSNEIRTCORHCGOYSAORSPHOGVD 57  
Db 206 ILAGFEALIHGVEIISISIGPLADY-----SSDSI-----SVGSFHAMR-----KGIL 250

Qy 58 XLCASGTYAPFTGMIV-----GOEKPYONKNAINNGVRISGRGFCVYKMPYI 105  
Db 251 TVASAGND--GPSSGVTVTHNEPWLIVVAGSIDRFTFSKIDLGKNSFSGWG--ISMESP 306

Qy 106 KPIKYKGPYKKG 117  
Db 307 KAKSY--PLVSG 316

## RESULT 14

E97836

probable periplasmic protein [imported] - Rickettsia conorii (strain Malish 7)

C:/Species: Rickettsia conorii

C:/Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001

C:/Accession: E97836

R/Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Rickettsia Science 293, 2093-2098, 2001

A:/Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.

A:/Reference number: A97700; M01D:21442074; PMID:11557893

A:/Accession: E97836

A:/Status: preliminary

A:/Molecule type: DNA

A:/Residues: 1-284 &lt;KUR&gt;

A:/Cross-references: GB:AE006914; PIDN:AL03631.1; PID:915620216; GSPDB:GN00173

C:/Genetics:

A:/Gene: RC1093

## Query Match

8.8%; Score 71; DB 2; Length 284;

Best Local Similarity 24.0%; Pred. No. 12;

Matches 24; Conservative 16; Mismatches 32; Indels 28; Gaps 5;

QY 53 HQGVDXLCSAGSTVAPFTGMIV-GQEKPYQNKNAINGVRISGRGFCVKMEYIKPIRY 110

Db 181 HSGIDLOAKKAPLYAAAGIVIKAAAPDYGNEVEIKH-----GRKFVTKYAHLEKMSV 235

QY 111 K-GPIKKGEKIGTLPLQKVPDQ-----SHVHIE 140

Db 236 KEGNKIKRKQ-----FTIGQSTGNATGEHLHFE 264

## RESULT 15

T05614

hypothetical protein F9D16.290 - Arabidopsis thaliana

C:/Species: Arabidopsis thaliana (mouse ear cress)

C:/Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 17-Mar-2000

C:/Accession: T05614

R/Bevan, M.; Wedler, H.; Wedler, E.; Wambut, R.; Hohnselt, J.; Mewes, H.W.; Mayer, K.F. submitted to the Protein Sequence Database, February 1999

A:/Reference number: 215419

A:/Accession: T05614

A:/Molecule type: DNA

A:/Residues: 1-444 &lt;BEV&gt;

A:/Cross-references: EMBL:AL035394

A:/Experimental source: cultivar Columbia; BAC clone F9D16

C:/Genetics:

A:/Map position: 4

A:/introns: 110/3; 151/1; 209/2; 233/1

A:/Note: F9D16.290

C:/Superfamily: polygalacturonase

## Query Match

8.7%; Score 70; DB 2; Length 444;

Best Local Similarity 25.7%; Pred. No. 23;

Matches 38; Conservative 14; Mismatches 54; Indels 42; Gaps 8;

QY 23 NICAAGSSNEIRTC-----RHGCGOYSAORSOPHOGVDXLCSTVYAPFTGMIVGO 77

Db 237 NVCI--EDSYISTGDLVAIKSGWDQYIAYG-RPSSNITTRITGS---SPFAGIATGS 290

QY 78 EKPYNKNAINGVRISGRGFCVKM-----FYIKPIK---YKGP1KKGKELG----- 121

Db 291 ETSGIGIKNIIEHITLSNMGVGVNKTINIGRGIGIKNIKISDYVDFTAKYGIK1AGDNGD 350

QY 122 -----TLPLQKVPDQ-----SHVHIE 141

Db 351 HPDENYNPALPVK-----GIHITN 371

Search completed: March 25, 2003, 17:22:09  
Job time : 31.5 secs





GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 25, 2003, 17:10:39 ; Search time 8.5 Seconds  
(without alignments)  
736.815 Million cell updates/sec

Title: US-09-868-953a-1

Perfect score: 804  
Sequence: 1 MESTKALLAGLISTALAGP.....GIQSHVHENCSSDPAYL 151

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	802	99.8	151	1	LCR2_HUMAN
2	698	86.8	151	1	LCR2_BOVIN
3	681	84.7	151	1	LCR2_MOUSE
4	429	53.4	326	1	MIM1_CHICK
5	76.5	9.5	475	1	YEB4_HAEN
6	73	9.1	710	1	DAS_PICAN
7	71	8.8	779	1	TRP3_THEVO
8	70.5	8.8	680	1	Z334_HUMAN
9	69.5	8.6	472	1	ZIM3_HUMAN
10	69	8.6	1330	1	XDH_RAT
11	68	8.5	184	1	Y763_CAEL
12	68	8.5	762	1	MUS2_AQUAE
13	67.5	8.4	355	1	COL1_ARATH
14	67.5	8.4	535	1	VG50_HSVSA
15	67.5	8.4	625	1	TR11_MOUSE
16	67	8.3	398	1	MOB1_XENLA
17	65.5	8.1	1241	1	NPHN_HUMAN
18	65	8.1	271	1	Y1BF_ECOLI
19	65	8.1	312	1	PGLR_AGRTU
20	65	8.1	317	1	IBP2_SHEEP
21	65	8.1	1388	1	RPOD_TOBAC
22	64.5	8.0	444	1	FA7_RABIT
23	64.5	8.0	571	1	SECD_MYCLE
24	64.5	8.0	2241	1	TEGU_HCVNA
25	64	8.0	593	1	MPE2_HUMAN
26	64	8.0	746	1	GYP7_YEAST
27	64	8.0	1477	1	KEIC_DROME
28	63.5	7.9	551	1	LEM2_RABIT
29	63	7.8	417	1	KCRU_HUMAN
30	63	7.8	490	1	HPOV_CLOAM
31	63	7.8	664	1	PM5A_STRMU
32	63	7.8	947	1	PM19_CHLPN
33	62.5	7.8	466	1	FA7_HUMAN

34	62.5	7.8	691	1	OAT6_HUMAN	O9y616 homo sapien
35	62	7.7	109	1	CYC_TERYX	P00079 tetranymena
36	62	7.7	221	1	GPI1_ZINOF	P82474 zingiber of
37	62	7.7	274	1	VPRP_MMTVB	P10271 mouse mamma
38	62	7.7	499	1	DHAS_CHICK	O93344 gallus galli
39	62	7.7	655	1	TR21_MOUSE	O9epus mus musculu
40	62	7.7	775	1	YTX1_XENLA	P14380 xenopus lae
41	62	7.7	948	1	HMDI_SCHEMA	P16237 schistosoma
42	62	7.7	1196	1	BXCN_CLOBO	P46081 clostridium
43	62	7.7	2109	1	RRLP_VERYH	P13615 vesicular s
44	62	7.7	3412	1	POLG_TREVS	P07720 t genome po
45	61.5	7.6	362	1	FIBP_ADE08	P36845 human adeno

## ALIGNMENTS

RESULT 1  
LCR2\_HUMAN STANDARD; PRT; 151 AA.  
ID O14960; O14565;  
AC 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Leukocyte cell-derived chemotaxin 2 precursor (hLECT2).  
GN LECT2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANT ILE-58.  
RC TISSUE=Liver;  
RX MEDLINE=98193133; PubMed=9524238;  
RA Yamagoe S., Mizuno S., Suzuki K.;  
RT "Molecular cloning of human and bovine LECT2 having a neutrophil  
RT chemotactic activity and its specific expression in the liver.";  
RL Blochim. Biophys. Acta 1396:105-113(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Periphera blood;  
RX MEDLINE=98207247; PubMed=9545637;  
RA Yamagoe S., Kameoka Y., Hashimoto K., Mizuno S., Suzuki K.;  
RT "Molecular cloning, structural characterization, and chromosomal  
RT mapping of the human LECT2 gene.";  
RL Genomics 48:324-329(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99048897; PubMed=9832057;  
RA Nagai H., Hamada T., Uchida T., Yamagoe S., Suzuki K.;  
RT "systemic expression of a newly recognized protein, LECT2, in the  
RT human body.";  
RL Pathol. Int. 48:882-886(1998).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,  
RA Kadner C., Miguel T., Miller C., Pittluck S., Pollard M., Rojestei H.,  
RA Subramanian S., Martin C.H.;  
RT Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Strong C., Biewald T., Tin-Wollam A., Duckels G.;  
RT Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SUBCELLULAR LOCATION.  
RX MEDLINE=97031488; PubMed=8877413;  
RA Yamagoe S., Yamakawa Y., Matsuo Y., Minowada J., Mizuno S., Suzuki K.;  
RT "Purification and primary amino acid sequence of a novel neutrophil  
RT chemotactic factor LECT2.";  
CC Immunol. Lett. 52:9-13(1996).  
CC -!- FUNCTION: Has a neutrophil chemotactic activity. Also a positive  
CC regulator of chondrocyte proliferation.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic and secreted.

CC -1- TISSUE SPECIFICITY: Highly expressed in adult and fetal liver and  
 CC weakly in testis. Not expressed in bone marrow.  
 CC -1- INDUCTION: By phytohemagglutinin (PHA).  
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: D63521; BAA23609.1; -;  
 DR EMBL: AB007546; BAA25669.1; -;  
 DR EMBL: AC004763; AAC17734.1; -;  
 DR EMBL: AC002428; AAB66905.1; -;  
 DR Genew; HGNC:6550; LECT2.  
 DR MIM; 602882; -;  
 DR Chemotaxis; Signal; Polymorphism.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 151 LEUKOCYTE CELL-DERIVED CHEMOTAXIN 2.  
 FT VARIANT 58 58 V->I.  
 SQ SEQUENCE 151 AA; 16376 MW; A4267F0A24B69631 CRC64;  
 /FTD-VAR-011386;  
 Query Match 99.8%; Score 802; DB 1; Length 151;  
 Best Local Similarity 99.3%; Pred. No. 3, 1e-77;  
 Matches 150; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MESTALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDYLC 60  
 DB 1 MESTALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDYLC 60  
 QY 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKMFYIKPIKYGPIKKGKEL 120  
 DB 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKMFYIKPIKYGPIKKGKEL 120  
 QY 121 GTLLPQKVPYPGIOSHVHIENCSDSPPTAYL 151  
 DB 121 GTLLPQKVPYPGIOSHVHIENCSDSPPTAYL 151  
 RESULT 2  
 LCT2\_BOVIN STANDARD; PRT; 151 AA.  
 ID LCT2\_BOVIN  
 AC 062644;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Leukocyte cell-derived chemotaxin 2 precursor (bLECT2) (Chondromodulin  
 DE II) (bChm-II).  
 GN LECT2.  
 OS Bos taurus (Bovine).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 CC NCB1\_TaxID=9913;  
 OX NCB1\_TaxID=9913;  
 RN RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RC MEDLINE=98193133; PubMed=9524238;  
 RA Yamagoe S., Mizuno S., Suzuki K.;  
 RT "Molecular cloning of human and bovine LECT2 having a neutrophil  
 RT chemotactic activity and its specific expression in the liver.";  
 RL Blochim. Biophys. Acta 1396:105-113(1998).  
 RN RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RC MEDLINE=99160594; PubMed=10050029;  
 RA Shukunami C., Kondo J., Takahashi K., Inoue H., Kamizono A.,  
 RA Hiraki Y.;  
 RT "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the

RT growth-promoting actions of bovine recombinant protein.";  
 RL J. Biochem. 125:436-442(1999).  
 RN RP SEQUENCE OF 19-151.  
 RP TISSUE=Epiphyseal cartilage;  
 RC MEDLINE=96394331; PubMed=8796437;  
 RA Hiraki Y., Inoue H., Kondo J., Kamizono A., Yoshitake Y.,  
 RA Shukunami C., Suzuki K.;  
 RT "A novel growth-promoting factor derived from fetal bovine cartilage,  
 RT chondromodulin II. Purification and amino acid sequence.";  
 RL J. Biol. Chem. 271:22657-22662(1996).  
 CC -1- FUNCTION: Has a neutrophil chemotactic activity. Also a positive  
 CC regulator of chondrocyte proliferation.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB001350; BAA25302.1; -;  
 DR EMBL: D89011; BAB18616.1; -;  
 DR Chemotaxis; Signal.  
 FT SIGNAL 1 18 LEUKOCYTE CELL-DERIVED CHEMOTAXIN 2.  
 FT CHAIN 19 151  
 FT SEQUENCE 151 AA; 16320 MW; 2A1D8F5B28A3B8 CRC64;  
 Query Match 86.8%; Score 698; DB 1; Length 151;  
 Best Local Similarity 86.1%; Pred. No. 2, 6e-66;  
 Matches 130; Conservative 7; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 MESTALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDYLC 60  
 DB 1 MESTALLAGLITSTALAGPWANICAGKSSNEIRTCDRHGGCGQYSAORSQRPBGVDYLC 60  
 QY 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKMFYIKPIKYGPIKKGKEL 120  
 DB 61 SAGSTVYAPFTGMIVGQEKPYONKNAIINGVRISGRGFCVKKMFYIKPIKYGPIKKGKEL 120  
 QY 121 GTLLPQKVPYPGIOSHVHIENCSDSPPTAYL 151  
 DB 121 GTLLPQKVPYPGIOSHVHIENCSDSPPTAYL 151  
 RESULT 3  
 LCT2\_MOUSE STANDARD; PRT; 151 AA.  
 ID LCT2\_MOUSE  
 AC 088803; O88804; O9GWN3; Q92337;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Leukocyte cell-derived chemotaxin 2 precursor (Chondromodulin II)  
 DE (Chm-II).  
 GN LECT2.  
 OS Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCB1\_TaxID=10090;  
 OX NCB1\_TaxID=10090;  
 RN RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANT VAL-129.  
 RC STRAIN=BALB/C; TISSUE=Liver;  
 RC MEDLINE=98382586; PubMed=9714793;  
 RA Yamagoe S., Watanabe T., Mizuno S., Suzuki K.;  
 RT "The mouse LECT2 gene: cloning of cDNA and genomic DNA, structural  
 RT characterization and chromosomal localization.";  
 RL Gene 216:171-178(1998).  
 RN RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=Swiss Webster / NIH Swiss; TISSUE=Embryo, and Liver;

```

RA MEDLINE=99160594; PubMed=10050029;
RX Shukunami C., Kondo J., Wakai H., Takahashi K., Inoue H., Kamizono A., Hiraki Y.; "Molecular cloning of mouse and bovine chondromodulin-II cDNAs and the growth-promoting actions of bovine recombinant protein."; RT J. Biochem. 125:436-442(1999).
RL - FUNCTION: Has a neutrophil chemotactic activity. Also a positive regulator of chondrocyte proliferation. CC
CC - SUBCELLULAR LOCATION: Secreted. CC
CC - ALTERNATIVE PRODUCTS: 2 isoforms; 1/Lect2 (shown here) and 2/LectQ2; are produced by alternative splicing. CC
CC - TISSUE SPECIFICITY: Highly expressed in liver and weakly in muscle. Not expressed in heart, brain, spleen, lung, skeletal CC
CC -- SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY. ----- CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announcement or send an email to licenses@isb-sib.ch). ----- DR
DR EMBL; AB009687; BAA33383.1; -
DR EMBL; AB009688; BAA33384.1; -
DR EMBL; AB009689; BAA33385.1; -
DR EMBL; AB009689; BAA33386.1; -
DR EMBL; AF035161; AAF13302.1; -.
KW GMD; MG1:1278342; Lect2.
KW Chemotaxis; Signal; Alternative splicing.
FT SIGNAL 18 BY SIMILARITY.
FT CHAIN 19 LEUCOCYTE CELL-DERIVED CHEMOTAXIN 2.
FT FT FCYVIFLYPIKIKGSIKKGEKLGTLLPLOKIYPGIOSH VENDORSDPTAVL -> ORNLHTTLNVTCTWDKIQIPR PTRLDCDPEFLH (IN ISOFORN LECT20).
FT VARSPPLIC 98 .
FT VARIANT 129 129 I->V.
FT SEQUENCE 151 AA; 16405 MW; 1BAF444046B7AE8E CRC64; SQ
Query Match      84.7%; Score 681; DB 1; Length 151;
Best Local Similarity 82.1% Pred. No.1.6e-64;
Matches 124; Conservative 12; Mismatches 15; Indels 0; Gaps 0;
OY   1 MFSTKALLAGLISTALAGPWANICAGKSNEIRFCRDHCSCGOYSQAORSOPRHGVDXLC 60
DB   1 MIPTLTLSALLSSALAGPANICASSSNEIRPCDSYGCGGYSAORTRQHHPGVYL 60
OY   61 SNGSYTPAPRTGMIVGEEKPYONKNAINNVGAISGRGFVCVMKYIPIKKGPKKEKL 120
DB   61 SGGSVVYAPTGGKITVGEEKPRNRNNAINDGIRLSRGFGCVKTIFYIKPIKKGSIKKEKL 120
OY   121 GTLLPLQKVYPGIGIOSHVIENCDSSDPAYLV 151
DB   121 GTLLPLQKITPGIOSHVHENCKDDSDPAYVL 151
RESULT 4
MIMI_CHICK STANDARD; PRT: 326 AA.
ID     IMI_1CHICK
AC    P08940;
DT    01-AUG-1990 (Rel. 15, Created)
DT    01-DEC-1992 (Rel. 24, last sequence update)
DT    15-JUL-1999 (Rel. 38, last annotation update)
DE    Myeloid proteIn-1 precursor (P33).
GN    MIM-1.
OS    Gallus gallus (Chicken).
OC    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC    Archosauroidae; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; OC    Gallus.
OX    NCBI_TaxId=9031;
RN    [1]
RP    SEQUENCE FROM N.A.
TX    MEDLINE=90090611; PubMed=2688896;
```

RA	Ness S.A.	Marknell A., Graf T.;
RA	"The v-myb oncogene product binds to and activates the promyelocyte-	
RT	specific m-m-1 gene.";	
RL	Cell 59:1115-1125(1989).	
RN	[2]	
RP	PARTIAL SEQUENCE, AND SEQUENCE OF 295-299 FROM N.A.	
RC	STRAIN-White leghorn;	
RX	MEDLINE=93011981; PubMed=1397316;	
RA	Yamada K., Tsunohya M., Mishima K., Shinoyama M.;	
RT	"p33, an endogenous target protein for arginine-specific ADP-	
RT	ribosyltransferase in chicken polymorphonuclear leukocytes, is highly	
RL	homologous to mim-1 protein (myb-induced myeloid protein-1).";	
FEBS Lett.	311:203-205(1992).	
-	FUNCTION: NOT KNOWN.	
-	SUBCELLULAR LOCATION: GRANULES OF PROMYELOCYTES.	
-	INDUCTION: BY THE MYB ONCOGENE.	
-	PTM: SUBSTRATE FOR ARGININE-SPECIFIC ADP-RIBOSYLTRANSFERASE.	
-	SIMILARITY: BELONGS TO THE LECT2 / MIM-1 FAMILY.	
-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	
DR	EMBL; M29448; AAA48954.1; -	
DR	EMBL; M29449; AAA48955.1; -	
DR	PIR; A33755; A33755.	
KW	PIR; S29229; S29229.	
DR	Repeat; Signal.	
FT	SIGNAL 1 23 PROBABLE.	
FT	CHAIN 24 326 MYELOID PROTEIN-1.	
FT	REPEAT 28 162 1.	
FT	REPEAT 177 312 2.	
FT	CONFLICT 297 297 I -> Y (IN REF. 1).	
SO	SEQUENCE 326 AA; 35636 MW; B8C4742EF0D3BAD3 CRC64;	
-----		
Query Match	53.4%; Score 429; DB 1; Length 326;	
Best Local Similarity	50.3%; Pred. No. 1.1e-37;	
Matches 80; Conservative 26; Mismatches 39; Indels 14; Gaps 2;		
QY	7 LLGLAGLISTALGP-----WANICGAKSSNEIRTCDRHCGGYSAQRSQR--P 52	
Dd	7 IALLSLVSTAFAKOMEVHPPOQGRHMNQJCSGNFNRJRGDCRGCGNYGASRGCGEK 66	
QY	53 HGGVDXLCASGSITVYAFPTGTIVGOEKPYOKNAINNGYRISGRGCYVMFYTKPKIKYG 112	
Dd	67 HGVGDVICDTGDSITVYAFPSQSLSGPIRFRRHNGMAIDDGQJSGSGCYVKLVCIHPTRYHG 126	
QY	113 PIKKGEKLGTLPLQKYYPGIOSHVIHENCSSDPATVL 151	
Dd	127 QIQKGQQLIGRMILPMQKVPEGLIVSHIHVENCDSDPTHL 165	
-----		
RESULT 5		
YEB_A_HAEIN	STANDARD;	PRT; 475 AA.
ID	YEB_A_HAEIN	
Ac	P44693;	
DT	01-NOV-1995 (Rel. 32, Created)	
DT	01-NOV-1995 (Rel. 32, Last sequence update)	
DT	15-JUN-2002 (Rel. 41, Last annotation update)	
DE	Hypothetical protein HI0409.	
GN	HI0409.	
OC	Haemophilus influenzae.	
OC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;	
OC	Haemophilus.	
OX	NCBI_TaxId=727;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=Rd / KW20 / ATCC 51907;	
RX	MEDLINE=95350630; PubMed=7542800;	

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Sherry R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- FUNCTION: COULD BE INVOLVED IN CELL WALL DEGRADATION OR  
 CC FORMATION (BY SIMILARITY).  
 CC -1- SIMILARITY: STRONG, TO E.COLI YEBA.  
 CC -1- SIMILARITY: TO STAPHYLOCOCCUS LYSOSTAPHIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL: U32724; AAC22068.1; -.  
 DR MEROPS: M37.UPW; -.  
 DR TIGR: H10409; -.  
 DR InterPro: IPR002886; Peptidase\_M37.  
 DR Pfam: PF01551; Peptidase\_M37.1.  
 DR Hypothetical protein: Cell wall; Complete proteome.  
 KW SEQUENCE 475 AA; 53255 MW; 143C10F92233939D CRC64;  
 SQ  
 Query Match 9.5%; Score 76.5; DB 1; Length 475;  
 Best Local Similarity 28.7%; Pred. No. 1.7;  
 Matches 27; Conservative 15; Mismatches 41; Indels 11; Gaps 4;  
 QY 51 RPHGVADLXCSAGSTYAPFTGMITVGOEKPYONKNAINGVRISGRGCVKMFYKRP--I 108  
 DB 346 RPHKGVADSVSOGTPVIADPADTGV--EKVAVYQAGAGRYVALRGRGQVTVYMHLSKSLV 403  
 QY 109 KYKPIKKGKXGLTLLPLQKVPYGIOS--HVHIE 140  
 DB 404 KAGQTVKKGRIAL-----SGNTGISTGPHDHYE 432

RESULT 6  
 DAS\_PICAN STANDARD: PRT; 710 AA.  
 AC P06834;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Dihydroxy-acetone synthase (EC 2.2.1.3) (DHAS) (formaldehyde  
 DE transketolase) (glycerone synthase).  
 GN DAS.  
 OS Pichia angusta (Yeast) (Hansenula polymorpha).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Pichia.  
 OX NCBI\_TaxID=4905;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 34438;  
 RX MEDLINE=85215670; PubMed=2987872;  
 RA JENOWICZ Z.A., Eckart M.R., Drawke C., Roggenkamp R.O.,  
 RA Hollenberg C.P., Maat J., Ledebor A.M., Visser C., Verrips C.T.;  
 RT "Cloning and characterization of the DAS gene encoding the major  
 RT methanol assimilatory enzyme from the methylotrophic yeast Hansenula  
 RT polymorpha.";  
 RL Nucleic Acids Res. 13:3043-3062(1985).  
 RN [2]  
 RP REVISIONS, SEQUENCE OF 667-710 FROM N.A.  
 RC STRAIN=ATCC 34438;

RX MEDLINE=93101130; PubMed=1465101;  
 RA Hansen H., Didion T., Thiemann A., Veenhuis M., Roggenkamp R.O.;  
 RT "Targeting sequences of the two major peroxisomal proteins in the  
 RT methylotrophic yeast Hansenula polymorpha.";  
 RL Mol. Gen. Genet. 235:269-278(1992).  
 CC -1- FUNCTION: THIS IS THE MAJOR METHANOL ASSIMILATORY ENZYME FROM  
 CC THE METHYLOTROPHIC HANSENULA POLYMORPHA.  
 CC -1- CATALYTIC ACTIVITY: D-xylulose 5-phosphate + formaldehyde =  
 CC glyceraldehyde 3-phosphate + glycetone.  
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal.  
 CC -1- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL: X02424; CAA26276.1; -.  
 DR PIR: A23009; XJHOFK.  
 DR HSP: P23254; TTRK.  
 DR InterPro: IPR000360; Transketolase.  
 DR Pfam: PF00456; transketolase; 1.  
 DR Pfam: PF02779; transketolase; 1.  
 DR Pfam: PF02780; transketolase; 1.  
 DR PROSITE: PS00801; TRANSKETOLASE\_1; 1.  
 DR PROSITE: PS00802; TRANSKETOLASE\_2; 1.  
 DR PROSITE: PS00342; MICROBODIES\_CTER; 1.  
 KW Methanol utilization; Peroxisome; transfease; Thiamine pyrophosphate.  
 FT SITE 708  
 SQ SEQUENCE 710 AA; 78842 MW; E424B6D37CC8B0B CRC64;  
 Query Match 9.1%; Score 73; DB 1; Length 710;  
 Best Local Similarity 24.1%; Pred. No. 6;  
 Matches 33; Conservative 21; Mismatches 49; Indels 34; Gaps 7;  
 QY 11 GLISTALGPMNINICAGKSSNEIRTCDRHCGOYSAQRPHOGVDLXCSAGSTYAPF 70  
 DB 525 GLFQKAVELPFSSILS-LSRNEVL-----GLASRAQGRRNAGYILEDAENAEVQI 575  
 QY 71 TGMTVGOEKPYONKNAINGVRISGRG-----CYKMTYPIPKKGP--KKGEL 120  
 DB 576 IG--VGAEMEFADKA-----KILGRKFRTRVLSIPTRLPEDQSIGRSLVRDGRGV 628  
 QY 121 GTLLPLQKVPYGIOSHV 137  
 DB 629 PTVV-----VDGHV 637

RESULT 7  
 TRF3\_THEVO STANDARD: PRT; 779 AA.  
 AC Q97AJ6;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tricorn protease interacting factor F3 (EC 3.4.11.-).  
 GN TRF3 OR TV0814 OR TVG0817891.  
 OS Thermoplasma volcanium.  
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 OC Thermoplasmataceae; Thermoplasma.  
 OX NCBI\_TaxID=50339;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GSSI / DSM 4299 / JCM 9571;  
 RX MEDLINE=20570466; PubMed=11121031;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
 RA Nunoshita T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic

RT sequence of Thermoplasma volcanium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 CC -1- FUNCTION: Proteases F1, F2 and F3 degrade oligopeptides produced  
 CC by tricorn (themselves probably produced by the proteasome),  
 CC yielding free amino acids (By similarity).  
 CC -1- COFACTOR: Binds one zinc ion (By similarity).  
 CC -1- SUBUNIT: Part of the tricorn proteolytic complex (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP000993; BAB59956.1; -;  
 DR MEROPS: M01.021; -;  
 DR InterPro: IPR001930; Ala-peptase.  
 DR InterPro: IPR000130; Zn.Mpeptase.  
 DR Pfam: PF01433; Peptidase.M1.1.  
 DR PRINTS: PR00756; ALADIPASE.  
 DR PROSITE: PS00142; ZINC\_PROTEASEF.1.  
 DR HydroLase: Metalloprotease; Aminopeptidase; Zinc; Complete proteome.  
 FT METAL 266 266 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 267 267 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 270 270 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT METAL 289 289 ZINC (CATALYTIC) (BY SIMILARITY).  
 FT ACT\_SITE 352 352 PROTON DONOR (POTENTIAL).  
 SQ SEQUENCE 779 AA; 88068 MW; 6BC227FE2ECECB201 CRC64;  
 Query Match 8.8%; Score 71; DB 1; Length 779;  
 Best local Similarity 26.4%; Pred. No. 11;  
 Matches 33; Conservative 21; Mismatches 59; Indels 12; Gaps 4;  
 QY 24 ICAGKSSNEIRTCORHGGGQYSAQSRPHQGVXLCASGTYVAPPTGMVGEKPYQN 83  
 DB 590 IAYGKASRLALLDEAYC-ETLAPRFSNFEQCPPELKSATATYALSTGDVKGWYKRS 648  
 QY 84 KNAINNGVR-ISSRG---FCVGMFYIKPIKYGPIKGEKLG-----TILPLQKVP 132  
 DB 649 LDRDEDVYKIIIGFEKLSSTDLVSVMIEKGIKQDMLSFYLSALETWAGREYIYSN 708  
 QY 133 IQSHV 137  
 DB 709 LENTV 713  
 RESULT 8  
 2334\_HUMAN STANDARD; PRT; 680 AA.  
 ID 2334\_HUMAN STANDARD; PRT; 680 AA.  
 AC 09HCL1; 09NYM4;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Zinc finger protein 334.  
 GN ZNF334.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxId=9606;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21638749; PubMed-11780052;  
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,  
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,  
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,  
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,  
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,  
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,  
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,

RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,  
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
 RA Gaffam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
 RA Levasialho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
 RA Marsh V.L., Martin S.L., McConnachie L.J., McIay K., McMurray A.A.,  
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
 RA Rice C.M., Ross M.T., Scott C.E., Sehna H.K., Showkeen R., Sims S.,  
 RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
 RA Swan R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
 RA Wilmberg L., Wray P.W., Hubbard T., Dudin R.M., Bentley D.R., Beck S.,  
 RA Rogers J.;  
 RT "The DNA sequence and comparative analysis of human chromosome 20.";  
 RL Nature 414:865-871(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Nagatsuma M., Hosofiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Maeno Y.,  
 RA Nimomiya K., Iwayanagi T.;  
 RT "NEDO human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Testis;  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-  
 CC FINGER PROTEINS.  
 CC -----  
 CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AL034424; CAC18856.1; -;  
 CC EMBL: AL354745; CAC03544.1; -;  
 CC EMBL: AK001331; BA091630.1; -;  
 CC EMBL: BC024177; AAH24177.1; -;  
 CC HSSP: P03001; ITP6.  
 CC Genew: HGNC:15806; ZNF334.  
 DR InterPro: IPR001909; KRAB.  
 DR InterPro: IPR000822; Znf.C2H2.  
 DR Pfam: PF00006; 2f.C2H2; 28.  
 DR Pfam: PF01352; KRAB; 2.  
 DR ProDom: PD000003; Znf.C2H2; 7.  
 DR SMART: SM00349; KRAB; 1.  
 DR SMART: SM00355; Znf.C2H2; 14.  
 DR PROSITE: PS50805; KRAB; 1.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 14.  
 DR PROSITE: PS0157; ZINC\_FINGER\_C2H2\_2; 14.  
 KW Transcription regulation; Zinc-finger; Metal-binding; Nuclear protein;  
 KW DNA-binding; Repeat.  
 FT DOMAIN 10 81 KRAB.  
 FT DOMAIN 237 678 ZINC\_FINGERS.  
 FT ZN\_FING 237 259 C2H2-TYPE.  
 FT ZN\_FING 265 287 C2H2-TYPE.  
 FT ZN\_FING 293 315 C2H2-TYPE.







DR EMBL: AE000730; AAC07247.1; -.  
 DR InterPro: IPR000302; KOW\_motif.  
 DR InterPro: IPR000432; Muts\_C.  
 DR InterPro: IPR002863; Muts\_N.  
 DR InterPro: IPR002625; Smr/Muts2.  
 DR Pfam: PF00467; KOW; 1.  
 DR Pfam: PF00488; Muts\_C; 1.  
 DR Pfam: PF01713; Smr; 1.  
 DR ProDom: PD001263; Muts\_C; 1.  
 DR SMART: SM00534; MutsC; 1.  
 DR SMART: SM00533; MutsC; 1.  
 DR SMART: SM00463; Smr; 1.  
 DR TIGRFAMs: TIGR01069; muts2; 1.  
 DR PROSITE: PS00486; DNA\_MISMATCH\_REPAIR\_2; FALSE\_NEG.  
 KW ATP-binding; DNA-binding; Complete proteome.  
 FT NP\_BIND 329 336 ATP (POTENTIAL).  
 SQ SEQUENCE 762 AA; 87538 MW; 3A7D977DC2EB01BC CRC64;

Query Match 8.5%; Score 68; DB 1; Length 762;  
 Best Local Similarity 21.8%; Pred. No. 22;  
 Matches 27; Conservative 19; Mismatches 40; Indels 38; Gaps 4;

QY 64 STVYAFETGIVGQEKRYOKNAINNGVRISGRF---CYKMFYIKPIKYKPIKGEK 119  
 DB 20 SYAHSPATEKIQNLKPYTKRKEKEIEISKAFEDIANVRLFFEDIREL--LKKAKL 77  
 QY 120 LGTL-----PLQVYVGIOSHVHEN--CDSPP 147  
 DB 78 QCAILGVEDIKILNINILKREIRVLSHVORLEPLKRYKKLYTFSPLENIITISIDP 137  
 QY 148 TAYL 151  
 DB 138 RGFV 141

## RESULT 13

COLI\_ARATH STANDARD; PRT; 355 AA.  
 AC 050055;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Zinc finger protein CONSTANS-Like 1.  
 GN COLI OR A55G15850 OR F1AF8\_230.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 CX NCBI\_TaxID=3702;

NP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Landsberg erecta; TISSUE=Aerial parts;  
 RA Puterbaugh J.J., Ledger S.E., Lee K., Robson F., Murphy G.,  
 RA Coupland G.;  
 RT "The flowering-time gene CONSTANS and homologue LIKE 1 exist  
 as a tandem repeat on chromosome 5 of Arabidopsis.";  
 RL (in) Plant Gene Register PGR97-077.

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Columbia;  
 RX MEDLINE=21016721; PubMed=11130714;  
 RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,  
 RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,  
 RA Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,  
 RA Nakazaki N., Nambu K., Okumura S., Shino S., Takeuchi C., Wada T.,  
 RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,  
 RA Huang B., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,  
 RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,  
 RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,  
 RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,  
 RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latrelle P.,  
 RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Stromwall C.,  
 RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,

RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vii D., Baker J.,  
 RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,  
 RA Marijnsen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,  
 RA Voicaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,  
 RA Entian K.-D., Terryn N., Hartley N., Bent E., Johnson S.,  
 RA Langham S.A., McCullagh B., Robben J., Grymoprez B., Zimmermann W.,  
 RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,  
 RA van Staveren M., Dirse W., Mooijman P., Klein Lankhorst R.,  
 RA Weitzenecker T., Bothe G., Rose M., Hauf J., Bernisier S., Hempel S.,  
 RA Feldpausch M., Lamberth S., Villarejo R., Gijlen J., Ardiles W.,  
 RA Bents O., Lemcke K., Kolesov G., Mayer K., Rüd S., Schoof H.,  
 RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.,  
 RA "Sequence and analysis of chromosome 5 of the plant Arabidopsis  
 thaliana".  
 RL Nature 408:823-826(2000).

RN [3]  
 RP CHARACTERIZATION.  
 RC STRAIN=cv. C24;  
 RX MEDLINE=21260187; PubMed=11359606;  
 RA Ledger S., Strayer C., Ashton F., Kay S.A., Puterbaugh J.,  
 RT "Analysis of the function of two circadian-regulated CONSTANS-Like  
 genes".  
 RL Plant J. 26:15-22(2001).  
 CC -1- FUNCTION: Putative transcription factor that may be involved in  
 the light input to the circadian clock but does not affect  
 flowering time.  
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LEAVES AND AT LOWER LEVELS  
 IN STEMS, FLOWERS AND STILIGUES. NOT DETECTED IN ROOTS.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT.  
 CC -1- INDUCTION: EXPRESSED WITH A CIRCADIAN RHYTHM SHOWING A PEAK  
 AT DAWN.  
 CC -1- SIMILARITY: BELONGS TO THE CONSTANS FAMILY.  
 CC -1- SIMILARITY: CONTAINS 2 B BOX-TYPE ZINC FINGERS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: Y10555; CAA71587.1; -.  
 CC EMBL: Y10556; CAA71588.1; -.  
 CC EMBL: AL391144; CAC01784.1; -.  
 DR InterPro: IPR000315; Znf\_Box.  
 DR InterPro: IPR002926; Znf\_Constans.  
 DR Pfam: PF00643; Zf-B\_Box; 2.  
 DR ProDom: PD007661; Znf\_Constans; 2.  
 DR SMART: SM00336; BBOX; 1.  
 DR PROSITE: PS50119; ZF\_BBOX; 2.  
 KW Transcription regulation; Zinc-finger; DNA-binding; Nuclear protein;  
 KW Repeat; Multigene family.  
 FT ZN\_FING 12 54 B BOX-TYPE 1.  
 FT ZN\_FING 55 97 B BOX-TYPE 2.  
 FT ZN\_FING 98 116 B BOX-TYPE 3.  
 SQ SEQUENCE 355 AA; 39491 MW; 412C237A9CD30309 CRC64;

Query Match 8.4%; Score 67.5; DB 1; Length 355;  
 Best Local Similarity 23.6%; Pred. No. 11;  
 Matches 29; Conservative 15; Mismatches 44; Indels 35; Gaps 4;

QY 20 PWANICAGKSSNEIRTCRRGCGYSAQSRPHQGVDXL-----CSAGSTV 66  
 DB 62 PAAPFCADAACTCTDSE---IHSANPLARRHQRVILPISEVSYSSTATTNHCSETTV 118  
 QY 67 YAPFTGIVGQEK-----PYONKAINNGVRISGRGCVKMFYIKPIKYKCP 113  
 DB 119 TDPENLVIGQEBEDEDAEASWLLPNSGKSGNN-----NGFSIGDEFILNLYVSSS 172  
 QY 114 IKK 116  
 DB 173 DRQ 175





```
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 625 TUMOR NECROSIS FACTOR RECEPTOR
FT SUPERFAMILY MEMBER 11A.
FT DOMAIN 31 214 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 215 234 POTENTIAL.
FT DOMAIN 235 625 CYTOPLASMIC (POTENTIAL).
FT REPEAT 35 69 TNFR-CYS 1.
FT REPEAT 72 113 TNFR-CYS 2.
FT REPEAT 115 152 TNFR-CYS 3.
FT REPEAT 155 195 TNFR-CYS 4.
FT DISULFID 35 47 BY SIMILARITY.
FT DISULFID 48 61 BY SIMILARITY.
FT DISULFID 51 69 BY SIMILARITY.
FT DISULFID 72 87 BY SIMILARITY.
FT DISULFID 93 113 BY SIMILARITY.
FT DISULFID 115 128 BY SIMILARITY.
FT DISULFID 134 152 BY SIMILARITY.
FT CARBOHYD 106 106 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 494 494 R -> K (IN REF. 2).
SQ SEQUENCE 625 AA; 66621 MW; F8C1872E99511D8E CRC64;

Query Match 8.48; Score 67.5; DB 1; Length 625;
Best Local Similarity 21.48; Pred. No. 20;
Matches 31; Conservative 11; Mismatches 40; Indels 63; Gaps 5;

QY 16 ALAGPWANI-----CAGKSSNEIRTCDRHGGCGYSAQRSQRPHQGVDXLCSA 62
DB 309 AAGGPWAEVRDSRFTLVSEVETQGLSRKIPTDEY-----TDRPSQFSTGSLLIQQ 362
QY 63 GSTVYAPFTGMIVGQEKPYQKNKAINNGVRISGRGFCVMEYIKPKYKGPDKKGLGT 122
DB 363 GSKSIPPF-----QE-----PLEVGENDS- 381
QY 123 LLPLQKVPYPGIOSHVHIENCDSSDP 147
DB 382 ---LSQCTGTSTVDSEGCDFTEP 403
```

Search completed: March 25, 2003, 17:20:05  
Job time : 16.5 secs



Q9DFJ1  
ID Q9DFJ1 PRELIMINARY; PRT; 156 AA.  
AC Q9DFJ1  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE Chemotaxin (Fragment).  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Bayne C.J., Gerwick L., Fujiki K., Nakao M., Yano T.;  
RT "Immune-relevant (including acute phase) genes identified in the  
RT livers of rainbow trout, Oncorhynchus mykiss, by means of suppression  
RT subtractive hybridization."  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF271114; AAG28030.1; -  
FT NON\_TER 156  
SQ SEQUENCE 156 AA; 17087 MW; B19067665FDA70DA CRC64;  
  
Query Match 46.0%; Score 370; DB 13; Length 156;  
Best Local Similarity 46.6%; Pred. No. 1.6e-32;  
Matches 68; Conservative 29; Mismatches 47; Indels 2; Gaps 1;  
  
QY 7 LLAGLSTALAGPANICAGKSSNEIRTCDRHGCGQYSAQR-SORPHQGVDXLCSAGSTV 66  
DB 10 VVLAVLSECEVMKFGQLCSNRRRTGDRWQGHGASRGRAHOGGLDIVNDGATV 69  
QY 67 YAPFTGMIVGQEKPYON--KNAINGVRISSGRGFCVKMFYIKPIKYGPIKKGKLGTL 124  
DB 70 YAPFDVLNGKVVYTPDKKAINDGINSGLGCLFVFKPKYSSVVKKGQRIGTML 129  
  
QY 125 PLQKYPGIQSHVHIENCSDSDPTAY 150  
DB 130 TMSVYPGITSRVHVQMCDSPTKF 155  
  
RESULT 3  
Q21241  
ID Q21241 PRELIMINARY; PRT; 472 AA.  
AC Q21241;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE Hypothetical 53.3 kDa protein.  
GN K05F1.5.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;  
RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Wohldmann P.;  
RT "The sequence of C. elegans cosmid K05F1.";  
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submission.";

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U29377; AAA68720.2; -  
KW Hypothetical protein.  
SQ SEQUENCE 472 AA; 53330 MW; D7A8D966A7190C2A CRC64;  
  
Query Match 17.7%; Score 142.5; DB 5; Length 472;  
Best Local Similarity 27.1%; Pred. No. 4.1e-07;  
Matches 45; Conservative 27; Mismatches 63; Indels 31; Gaps 9;  
  
QY 7 LLAGL-ISTALAGPANICAGKSSNEIRTC-DRHG-CGOYSAQR-SORPHQGVDXLCSA 62  
DB 146 LILIGIFIGGOTLGMKVCYGNNAKNDRCPTDGCNGYHRTSGEIDGVDVVRCHL 205  
QY 63 GSTVYAPFTGMIVGQEKPY--QKNAINNGVRISG----RGFCVKMFYIKPIKYGPIK 115  
DB 206 GEPIYAPIEGEMYEFW-RPYGGKREKSCADQGVRIEGTGWQGVAVHSSVKLSFFGGHVE 264  
QY 116 KGEKLGTLPL---QKYPGIGTQSHV-----HIENC 142  
DB 265 AGDEIGBALNRYCFNDRGQNDVPEHVEIRLYKEGRLLIDPTHLLQNC 310  
  
RESULT 4  
O01719  
ID O01719 PRELIMINARY; PRT; 689 AA.  
AC O01719;  
DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Myb-related transcription factor (Fragment).  
OS Strongylocentrotus purpuratus (Purple sea urchin).  
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
OC Echinioidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;  
OC Strongylocentrotus.  
OX NCBI\_TaxID=7668;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9808684; PubMed=9428408;  
RA Coffman J.A., Kirchhamer C.V., Harrington M.G., Davidson E.H.;  
RT "SpMyb functions as an intramodular repressor to regulate spatial  
RT expression of Cyilia in sea urchin embryos."  
RL Development 124:4717-4727(1997).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 3 MYB-LIKE DOMAINS.  
DR EMBL: U96090; AAC47807.1; -  
DR HSSP; P06876; IMBK.  
DR InterPro; IPR001005; Myb\_DNA\_binding.  
DR Pfam; PF00249; myb\_DNA\_binding; 3.  
DR SMART; SM00395; SANT; 3.  
DR PROSITE; PS00037; MYB\_1; UNKNOWN\_3.  
DR PROSITE; PS00334; MYB\_2; 3.  
DR PROSITE; PS50090; MYB\_3; 3.  
KW DNA-binding; Nuclear protein.  
FT NON\_TER 1  
SQ SEQUENCE 689 AA; 77241 MW; 0FA5A8D7AF8766C6 CRC64;

Query Match 10.9%; Score 87.5; DB 5; Length 689;  
Best Local Similarity 22.4%; Pred. No. 0.65; Indels 57; Gaps 9;  
Matches 36; Conservative 22; Mismatches 46;  
  
QY 17 LAGPWANIC---AGKSSNEIR-----TCDR--HGCGYSAQR-SOR-PH---QGVDXILCS 61  
DB 163 LGNRWAEIAKLLPGRTDNDNAIKHNWSTMRKRVETCNPTSKTTKTPHYTHNDNQKPCS 222  
QY 62 AGSTVYAPFTGMIVGQEKPYQKNAINNGVRISGRFCVKMFYIKPIKYGPIKKGKLG 121  
DB 223 SSSKIYTP-----DSDFMANSIRDALMRGQGRV-----VR 255  
QY 122 TLLPL-----QKYPGI---QSHVHIENCSDSDP 147  
DB 256 TLYPMGHDTLQDEGEKSGKVPKGVKTPQKWLIMDNCGEISP 296

```

RESULT 5
Q96ZEB PRELIMINARY; PRT; 272 AA.
AC Q96ZEB;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein S71885.
GN ST1885.
OS Sulfolobus tokodaii.
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
OC Sulfolobus.
OX NCBI_TaxID=111955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JCM 10545 / 7;
RX PubMed=11572479;
RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
RA Oshima T., Kikuchi H.;
RT "Complete genome sequence of an aerobic thermoacidophilic
RT Crenarchaeon, Sulfolobus tokodaii strain".
RL DNA Res. 8:123-140(2001).
DR EMBL; AF000988; BAB66977.1; -.
DR InterPro; IPR000515; BPD_transp.
DR PROSITE; PS00402; BPD_TRANSF.INN.MEMBR; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 272 AA; 30871 MW; 810C97413773BF05 CRC64;

Query Match 10.2%; Score 82; DB 17; Length 272;
Best Local Similarity 29.7%; Pred. No. 0.88;
Matches 27; Conservative 18; Mismatches 34; Indels 12; Gaps 5;

QY 67 YAPFTGMI-----VGQEKPYQNKNAINGVRISGRFCVKNFYIKPIKYKG-PIKKE 118
Db 36 YSPFGIIIEKIEKVKIGKPNKY-TKIDYDVVYINSNGRIKVLHVEPYLEEGSEIKE 94
QY 119 KLGTLPLQKVY-PGIIQSHVHIENCDSSDPT 148
Db 95 KIGKF--LESFYTAGDKHAIEGITFEPS 123

RESULT 6
Q9PGX9 PRELIMINARY; PRT; 417 AA.
AC Q9PGX9;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical protein Xf0167.
GN Xf0167.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,

```

```

RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.B.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.B., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
DR EMBL; AE003870; AAF92980.1; -.
DR InterPro; IPR002886; Peptidase_M37.
DR Pfam; PF01551; Peptidase_M37; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 417 AA; 46554 MW; 5483743CD61061FD CRC64;

Query Match 10.0%; Score 80.5; DB 16; Length 417;
Best Local Similarity 25.0%; Pred. No. 2.1;
Matches 29; Conservative 24; Mismatches 42; Indels 21; Gaps 7;

QY 34 RTCDRHCGGQYSAORSQPHQVDXLCSGSIVYAFPTGMI--VGQEKPYQNKNAINGV 91
Db 249 RLSSTFLRLRRPILGTMRMHKGVDAATGTPMAAGDARVQFQIGQYGVVILNH-- 306
QY 92 RISGRFCVKNFYIKPIKYKGPIKKEKL--GFLPLQKVYPGIGS----HVITE 140
Db 307 --GKGY--RTLYGHMSRF-GKIKAGQKINQGTIVG----YVGMTGLATGPHLHYE 352

RESULT 7
Q9PTR2 PRELIMINARY; PRT; 498 AA.
AC Q9PTR2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Cytochrome P450 2P2.
GN CYP2P2.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Fundulidae; Fundulus.
OX NCBI_TaxID=8078;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Stegeman J.J.;
RT "Diversity of vertebrate cytochrome P450 2 family genes: Cloning of
RT three new genes and phylogenetic analysis of Cyp2 gene subfamilies in
RT Fundulus heteroclitus";
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Oleksiak M.F., Wu S., Parker C., Zeldin D.C., Stegeman J.J.;
RT "Cloning and expression of a novel teleost cytochrome P450, CYP2P3:
RT conservation of arachidonic acid epoxidase/19-Hydroxylase.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
DR EMBL; AF117342; AAF21999.1; -.
DR HSSP; P00179; 1DT6.
DR InterPro; IPR001128; Cytochrome_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.

```

```
KW Heme: Monooxygenase; Oxidoreductase.
SQ SEQUENCE 498 AA; 57340 MW; 1FF4ED84B8D610C9 CRC64;

Query Match 9.7%; Score 78; DB 13; Length 498;
Best Local Similarity 21.7%; Pred. No. 4.9;
Matches 40; Conservative 28; Mismatches 58; Indels 58; Gaps 10;

QY 1 MFSTKALLAGLISTALAGPW-----ANICAGKSSNEIRTCDRHGGQYSAQRS 49
DB 117 VFNKGIVMSN-----GYPNKVRFFALHNRFLGKKTME-----RYIQOEC 160

QY 50 Q-----RPHQGVDXLCSAGSTVYAPFTGMIVGQEKPYQNK-----NAINGVR 92
DB 161 QYLNEVEFVDOQKPFSGGTLINNAVSN-----ICCLVGNRFYDDREYHTYLTIDNMNELLR 217

QY 93 ISRGFCVCFMFIPIKVK---GPTKKGKGLTLLPLQKVYPGQIOSHV---HIENCDSSDP 147
DB 218 LQG-GFVQVYVMFPMKWLPGPHKK-----IFIHLQIILDFLEIRKEHRELDPSSP 271

QY 148 TAYL 151
DB 272 RDYI 275

RESULT 8
Q9IED5 PRELIMINARY; PRT; 538 AA.
AC Q9IED5;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Env polypeptide (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RC STRAIN=YBF51;
RA Raoult P., Robertson D., Damond F., Sousquiere S., Maucelere P.,
RA Depienne C., Brun-Vezinet F., Dormont D.;
RT "Hiv-1 group O phylogenetic analysis of C2-gp41 region.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ133072; CAB96233.1;
DR InterPro; IPR000328; Env-GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER 1
FT NON_TER 538
SQ SEQUENCE 538 AA; 60777 MW; B3C9E66A233FEF1D CRC64;

Query Match 9.6%; Score 77.5; DB 15; Length 538;
Best Local Similarity 26.9%; Pred. No. 6;
Matches 36; Conservative 13; Mismatches 42; Indels 43; Gaps 6;

QY 25 CAGKSSN-EIRTCDRHGGQYSAQRSRPHQGVDXLCSAGSTVYAP-----FT 71
DB 175 CEGKSCNFKERT-----NNNTQIPCKLKQVRSWMKGGSLYAPPIRGNLCRSNIT 227

QY 72 GMIVGQEKPYQNKNAINGVRISG-----RGFCVMKMYIKPIKRGPI--- 114
DB 228 GMILQDKPW-NRNDTNTFRPIGGNMKDIWRTELLRYKVVRIKPFVAPTRIARPVIGS 286

QY 115 -----KKGKGLT 123
DB 287 GTHREKRAVGLGL 300

RESULT 9
Q9HS60 PRELIMINARY; PRT; 240 AA.
ID Q9HS60
```

RP	SEQUENCE FROM N.A.	Query Match	Best Local Similarity	Matches	41: Conservative	25: Mismatches	56: Indels	62: Gaps
RC	STRAIN=UAB CTIP;	9.2%;	Score 74;	DB 16;	Length 750;			
RX	MEDLINE=21267165; PubMed=11353084;	22.3%;	Pred. No. 21;					
RA	Chambaud I., Heilig R., Ferris S., Barbe V., Samson D., Galisson F.,							
RA	Moszer I., Dybvig K., Wroblewski H., Viari A., Rocha E.P.C.,							
RA	Blanchard A.;							
RT	"The complete genome sequence of the murine respiratory pathogen							
RT	Mycoplasma pulmonis";							
RL	Nucleic Acids Res. 29:2145-2153(2001).							
RL	EMBL; AL445564; CAC13651.1; -.							
DR	MypuList; MYP0_4780; -.							
KW	Complete proteome.							
QY	SEQUENCE 750 AA; 86325 MW; 37FF06FE161602D5 CRC64;							
Db								
QY	12 LISTALAGPWANICAGKSSNEIRTCDR--HCGGQYSAORSORPHOGVDXLCSSAGSTVYAP 69							
Db	456 LQSPVSGWLD-----DRSYSTSKASDKNTKDIHLGEDILIEQNKEVIAP 502							
QY	70 FTGMIVGQ---EKPYQ-----NKNAINNGVRISGRGFCVKMF-- 103							
Db	503 PDGKIIASYAPSPYQAYGLGVITVLEVMKKDLVGQIDQSDVIDNOLAETDRIYIAFMHLEN 562							
QY	104 -----YTKPIKYK-----PIKKEKLCGLTLPLQKVPYGIQSHVHIE-NCD 143							
Db	563 PSYLENYGKLVQVQSSSTAIEITPATPKTVKKEVIG-LVGEFRKNGGMPPHVVHIEVSLG 621							
QY	144 SSDP 147							
Db	622 STNP 625							
RESULT 13								
QY	Q8XVJ5 PRELIMINARY; PRT; 321 AA.							
AC	Q8XVJ5;							
DT	01-MAR-2002 (TReMBLrel. 20, Created)							
DT	01-MAR-2002 (TReMBLrel. 20, Last sequence update)							
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)							
DE	Putative transmembrane protein.							
GN	RSC2835 OR RS00270.							
OS	Ralstonia solanacearum (Pseudomonas solanacearum).							
OC	Bacteria; Proteobacteria; beta subdivision; Ralstonia group;							
OC	Ralstonia.							
RN	NCBI_TaxID=305;							
OX	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=GM11000;							
RX	MEDLINE=21681879; PubMed=11823852;							
RA	Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,							
RA	Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,							
RA	Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,							
RA	Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,							
RA	Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,							
RA	Weissenbach J., Boucher C.A.;							
RT	"Genome sequence of the plant pathogen Ralstonia solanacearum.";							
RL	Nature 415:497-502(2002).							
RL	EMBL; AL646072; CAD16542.1; -.							
DR	InterPro: IPR002886; Peptidase_M37.							
DR	Pfam; PF01551; Peptidase_M37; I.							
KW	Complete proteome.							
QY	SEQUENCE 321 AA; 33989 MW; 85011B16D1D82393 CRC64;							
Db								
Query Match	9.1%;	Score 73.5;	DB 16;	Length 321;				
Best Local Similarity	23.8%;	Pred. No. 9;						
Matches	30: Conservative	26: Mismatches	49: Indels	21: Gaps				
QY	37 DRHGGC----QYSAORSRPHOGVDXLCSSAGSTVYAPFTGMIVGQPKYQKNKAIN----- 88							
Db	174 DSSGFGVRIDPPTGRTO--HDCVDFRPGVGTPTVAAGGVVAAASEYHHEHNMIDIHG 231							

